Author:
Dadeh, Tarmidi H D

Title:
Exploring school climate and Indonesian Year 8 students' mathematics learning outcomes
a mixed method study

EXPLORING SCHOOL CLIMATE AND INDONESIAN YEAR 8 STUDENTS' MATHEMATICS LEARNING OUTCOMES: A MIXED METHOD STUDY

by

Tarmidi Hasballah Dadeh

A dissertation submitted to the University of Bristol in accordance with the requirements for the award of the degree of Doctor of Philosophy (PhD) in the Faculty of Social Science and Law.

SCHOOL OF EDUCATION

2021

ABSTRACT

The present study aims to explore the extent that school and classroom climate are related to students' learning outcomes in mathematics. The study incorporates both academic and non-academic achievements in Indonesia. By using a mixed-method approach, the study also seeks explanations for the observed differences in school climate between schools, based on the schools' academic performance and types of school (religious vs non-religious schools).

In the first phase, the author used a quantitative approach to examine the relationship between school and classroom climate with student mathematics assessment outcomes. The Indonesian TIMSS 2011 data of 5795 students, nested within 174 classrooms, and nested within 153 schools were analysed using three-level multilevel models (student, classroom, school). The results showed two school and classroom climate factors that had a significant positive effect on academic achievement, after controlling for student, teacher, and school characteristics: (1) student engagement in math lesson; and (2) school discipline. However, student safety had a significant negative relationship with academic achievement. On the other hand, equivalent analyses using student mathematics self-belief outcomes suggested five school climate factors of academic achievement: (1) student engagement with their school, (2) teacher safety, (3) teacher confidence in teaching math, (4) school physical resources, and (5) student safety. The other two factors that harmed academic achievement were: (1) teacher to teacher interaction, and (2) student engagement in mathematical lessons. However, these findings are tentative considering that the overall goodness of fit of the models was relatively weak (13-26 per cent).

In the second phase, the author employed a qualitative approach to examine whether the profile of the schools and classroom climate varied in schools that differed in terms of student academic performance and religious type. Four secondary schools were selected as illustrative cases. One high and one low academic performing school for each the general (non-religious school) and *madrasah* (religious school) were selected. Focus group interviews were conducted to four students and teachers. Individual interviews were conducted to respectively four headteachers and policymakers. The data was analysed using an analytical framework drawn from the literature. The study identified similarities and differences in school climate factors across the four sampled schools.

The study highlights the value of using the effectiveness evaluation dimensions of the Dynamic Model of Educational Effectiveness (DMEE) proposed by Creemers and Kyriakides (2008) (frequency, focus, stage, quality and differentiation) in exploring and evaluating school climate. By using the effectiveness evaluation dimensions, richer explanation about the differences between school climates of the high and low performing schools was obtained. The qualitative analysis also supports the findings in phase one, particularly those differences in school climate between the high and low performing schools. The results also expanded the framework by including two emerging factors: education orientation and culture.

Although the findings are tentative and exploratory, the research is original in particular. The study analyses and reports the relationship between school climate factors and broader student learning outcomes. The study also adopts a newly developed school effectiveness conceptual framework for Indonesia and other similar contexts. The study thus contributes to the international school effectiveness knowledge base.

Keywords: school effectiveness, school climate, TIMSS, math achievement, self-beliefs, Indonesia, multilevel modelling, mixed-method.

ACKNOWLEDGEMENT

Praise ALLAH, the Lord of the whole universe, for giving me blessing, strength, ability, and toughness for the completion of this dissertation.

I would like to convey my genuine appreciation to my supervisor Professor Sally Thomas. I am grateful for her advice, supportive comments, suggestions and support in finalising this study. Her commitment, motivation and encouragement have been significant for the accomplishment of this dissertation. Thank you to Professor Guoxing Yu, my second supervisor, for considering my first email that marked the beginning of my doctoral journey at the University of Bristol.

My recognition goes to the Government of the Republic of Indonesia. Particularly the Ministry of Research, Technology and Higher Education for their financial support. My gratitude goes to the former Dean of Faculty of Psychology, University of Sumatera Utara, Professor Irmawati, as well as the current Dean Dr Zulkarnain for letting me leave my duties as staff members to pursue a doctoral degree. I would like to appreciate Dr Angelina, the PhD programme Director, as well as the staff members of the School of Education, who have all contributed to my development as a researcher.

Finally, I would like to express my warmest gratitude to my beloved wife, Lolly Benro. Her understanding, love and support made this journey more relaxed. To my children, my first son Fatahillah, my only daughter Dara Hafiyya, my little sons Dawud Hamad and Dayyan Harun (the latter three kids were born during my study; two in Bristol and one in Medan), I am genuinely grateful for being your father. I love all of you. My warm-hearted appreciation also goes to my colleagues and friends who had supported me both morally and financially.

AUTHOR’S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programs and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: Tarmidi Dadeh DATE: 24 March 2021

DEDICATION

This dissertation is dedicated to my father Mr Hasballah Dadeh, who passed away during my PhD journey on 9th April 2018.

To my children:

Fatahillah Patilodi

Dara Hafiyya Patilodi

Dawud Hamad Patilodi

Dayyan Harun Patilodi

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENT iv

AUTHOR'S DECLARATION v

DEDICATION vi

FIGURES xvi

TABLES xvii

Chapter 1: Introduction 1

1.1 Introduction to the chapter 1

1.2 Research problems and backgrounds 1

1.3 Rationale for research 4

1.3.1 The importance of linking school climate with a range of learning outcomes as a means to promote school effectiveness 4

1.3.2 Indonesia Education Act and creating a positive school climate: The national context rationale 8

1.3.3 Personal rationale 11

1.4 Research aims and objectives 12

1.5 Research questions 14

1.6 Overview of the theoretical framework 15

1.7 Overview of the methodological framework 16

1.8 Brief description of the thesis by chapters 17

Chapter 2: The Indonesian context 19

2.1 Introduction to the chapter 19

2.2 Indonesia geographical location and culture 19

2.2.1 Geographic location and people 19

2.2.2 The Indonesian culture 21

2.3 Indonesia National Education System (INES) in brief 26

2.3.1 Basic education 28

2.3.2 Secondary education 29

2.3.3 Higher education 29

2.4 Challenges in Indonesia education system: existing strengths and weaknesses 29

2.4.1 Low learning outcomes compared to other countries 30

2.4.2 Quality of assessment 31

2.4.3 Quality of assessment 32

2.4.4 Constantly changed curriculum 33

2.4.5 The teaching-learning process: teacher-centred 35

2.4.6 Decentralisation: large quality gap across the nation 36

2.4.7 Improving teacher quality 38

2.5 Indonesia's *Madrasah* in brief 39

2.6 Chapter summary 41

Chapter 3: Literature review 43

3.1 Introduction to the chapter 43

3.2 A brief history of school effectiveness research (SER) 43

3.2.1 First phase 44

3.2.2 Critics on SER 48

3.2.3 Choosing among the various models of SER 50

3.2.4 Dynamic models of educational effectiveness (DMEE) 51

3.2.5 Limitations in using DMEE in the study 59

3.3 School climate 60

3.3.1 Significance of school climate in SER and its critiques 60

3.3.2 School climate: a challenging definition 63

3.3.3 School culture and school climate 66

3.3.4 Classroom climate and school climate 69

3.3.5 Issues in school climate research 70

3.3.6 School climate factors 73

3.3.7 Putting it all together 81

3.4 Learning outcomes 83

3.4.1 The non-cognitive aspect of learning 85

3.4.2 Academic self-beliefs 86

3.4.3 Relationship between achievement and academic self-concept 88

3.4.4 The importance of school as the social context in self-beliefs development 89

3.4.5 Learning outcomes measure in international study and some critics 90

3.4.6 Choosing TIMSS over PISA....94
3.5 Chapter summary....94
Chapter 4: Methodology....96
4.1 Introduction to the chapter....96
4.2 The philosophical approach to research design....97
4.2.1 Paradigm polarity: Positivism vs interpretivism....97
4.2.2 Chosen paradigm: Pragmatism....99
4.3 Methodological approach....103
4.3.1 Mixed-method research (MMR)....103
4.4 Quantitative phase....107
4.4.1 Secondary data analysis....107
4.4.2 Data preparation....109
4.4.3 TIMSS data....115
4.4.4 Re-analysis of climate scales: Rasch Model....117
4.4.5 Re-analysis of students' self-beliefs scales: Rasch Model....122
4.4.6 Variable used in the study....124
4.4.7 TIMSS 2011 data analysis plan: Multilevel model (MLM)....131
4.4.8 Limitations of MLM....134
4.5 Qualitative phase....134
4.5.1 Using illustrative case study schools....135
4.5.2 Data collection techniques and procedures....135

4.5.3 Qualitative data analysis 140

4.5.4 *A-priori* themes: 5 school climate factors combined with effectiveness evaluation dimensions 141

4.5.5 Data-driven categories 143

4.5.6 Strengths and limitations of the qualitative approach adopted 144

4.5.7 Trustworthiness 145

4.6 Ethical issues 145

4.6.1 Data protection 145

4.6.2 Informed consent 146

4.6.3 Researcher's position 146

4.6.4 Other moral concerns 147

4.7 Chapter summary 147

Chapter 5: Quantitative Findings 149

5.1 Introduction to the chapter 149

5.2 RQ1: Difference in mathematics and self-beliefs 149

5.3 RQ1: Difference in mathematics and self-belief controlling for school stakeholders' characteristics 151

5.3.1 Student characteristics 154

5.3.2 Teacher characteristics 156

5.3.3 School characteristics 161

5.4 RQ1.3: School climate factors that significantly explain the variance between school and classroom performance 168

5.4.1 Significant school climate factors before adjusting the characteristics of the student, teacher, and school 168

5.4.2 The impact of school climate factors after adjusting the characteristics of the student, teacher, and school 174

5.5 The achievement gap between general school and *madrasah* 176

5. 2. Chapter summary 181

Chapter 6: Qualitative findings 185

6.1 Introduction to the chapter 185

6.2 Stakeholders' perception of school climate 185

6.2.1 Similarities across all schools 185

6.2.2 Differences between high and low performing schools 195

6.2.3 Differences between *madrasah* and general school 211

6.3 New potential factors that emerged from the data analysis relevant to highlight differences between the high and low performed schools 212

6.3.1 The orientation of education 213

6.3.2 School culture 215

6.4 Chapter summary 216

Chapter 7: Discussion and Conclusions 218

7.1 Introduction to the chapter 218

7.2 Discussion and implications of key findings 218

7.2.1 Discussion of key findings of RQl: What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences? 218

7.2.2 Implications of key findings RQ1 229

7.2.3 Discussion of key findings of RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)? 230

7.2.4 Implications of key findings RQ2 235

7.3 Overall conclusions: linking RQ1 and RQ2 findings and implications 236

7.3.1 Conclusion 1: School climate relationship with student outcome learning were explained in the Indonesian context. 236

7.3.2 Conclusion 2: The use of effectiveness dimension of the dynamic model of educational effectiveness in assessing school and classroom climate were useful and give a more informative explanation. 237

7.3.3 Conclusion 3: Potential additional factors for evaluating school climate in the Indonesian context 237

7.4 Strengths and limitations of the study 239

7.4.1 Strengths of the study 239

7.4.2 Limitations of the study 240

7.5 Further research 242

7.5.1 Using longitudinal data or trend data analysis. 242
7.5.2 Using an improved measurement of school and classroom climate factors .. 243
7.5.3 Using different academic achievement 243
7.5.4 Looking more in-depth at the different aspect of differences between schools 244
7.6 Recommendations 244
7.7 Concluding statement 246
References 248
Appendices 313
Appendix 1: Ethics approval 313
Appendix 2: Letter of Acceses 317
Appendix 3: Peer review ethics form 320
Appendix 4: Consent form 324
Appendix 5: Participant Information sheet 327
Appendix 6: TIMSS variables used in the study 329
School climate 329
Self-beliefs 332
Appendix 7: example of school climate scale re-construction (winsteps) 333
Item fit 333
Item-person map 334
Unidimentionality 335

Appendix 8: Qualitative data gathering guidance.....336

Interview.....336

Interview Guidance.....336

Focus group interview (FGI).....338

FGI Guidance.....338

Appendix 9: New themes.....341

FIGURES

Figure 2-1. Map of Indonesia.....19

Figure 2-2. Indonesian population growth.....20

Figure 2-3. Religion in Indonesia 20

Figure 2-4. Indonesian culture index 22

Figure 2-5. Indonesian student achievement progress.....31

Figure 2-6. Reading proficiency among middle- and lower-income countries (from PISA 2000-2002).....32

Figure 3-1. Input process output model (Scheerens et al., 2003) 46

Figure 4-1: The sequential explanatory mixed-method procedure (adopted from Ivankova et al. (2006).....106

Figure 4-2: Q-Q plot 112

Figure 4-3: School climate effectiveness model.....142

TABLES

Table 2-1: Basic and secondary schooling system in Indonesia 27

Table 2-2: Islamic school forms in Indonesia 40

Table 3-1: School climate factors comparison 82

Table 4-1: Summary statistics of mathematic achievement data 109

Table 4-2: Explanation of parameter-level mean-square fit measurements 119

Table 4-3: Fit statistics for school climate factor scales 122

Table 4-4: Fit statistic of self-beliefs scales 124

Table 4-5: Summary statistics of mathematic achievement data 124

Table 4-6: Descriptive statistics of Indonesian students' self-beliefs 124

Table 4-7: Statistics descriptive for school climate factor scales 125

Table 4-8: Summary statistic of student's gender 126

Table 4-9 Summary statistic of student's gender 127

Table 4-10: Summary statistic of the frequency using Bahasa Indonesia 128

Table 4-11: Teacher gender 128

Table 4-12: Teacher age 128

Table 4-13: Teacher educational background 129

Table 4-14: Teaching experiences 129

Table 4-15: Summary statistic of average student's SES at the school level 130

Table 4-16: School types 130

Table 4-17: School location 131

Table 4-18: Research questions and multilevel models 133

Table 4-19: Participant in each school 138

Table 4-20: School classification.....139
Table 4-21: Profile of participated schools.....140
Table 5-1: Null Model for mathematics, math self-concept, and math self-efficacy 152
Table 5-2: Model 1 with student's backgrounds variables.....152
Table 5-3: Model 2 teacher's backgrounds as explanatory variables.....159
Table 5-4: Model 3 school's backgrounds as explanatory variables.....162
Table 5-5: Model 4 student, teacher, and school's backgrounds as explanatory variables ...163
Table 5-6: Model 5 school climate factors 170
Table 5-7: Model 6 Final Model.....178

Chapter 1: Introduction

1.1 Introduction to the chapter

The present research extends our knowledge about how school and classroom climate are related to students learning outcomes. It offers explanations that inform, enhance and provide recommendations on education policies and practices in Indonesia and the international context. The present introductory chapter will identify the research problem. It will also present the author's academic, local, and personal reasons to conduct the research. Next, the author describes six study objectives to clarify the nature of the current study. Subsequently, the author proposes two general research questions before summarising the rationale for the chosen research methods.

1.2 Research problems and backgrounds

A school is a complex, multilevel, and multifaceted institution. It is a designed and structured environment to stimulate children's behavioural and intellectual development. It is embedded in a broader social setting but is autonomous in many respects (Jones, 1995). Consequently, each school has differing characteristics, problems, and challenges (Freiberg, 1999, p. 3). In other words, every school has their personality, context and culture. The personality of a school is attributed to its school climate (Hoy, 2012). Hoy and his colleagues stated, “the climate may roughly be conceived as the personality of the school; that is, personality as to an individual as the climate is to an organisation” (Hoy et al., 1991, p. 134). Cohen et al. (2009) stated that “school climate refers to the character and quality of school life” (p. 100) and constructed based on patterns of school members' feelings and experiences of their school life.

A crucial problem in education is understanding the relationship between school and classroom climate with student learning outcomes. The present study seeks to address this issue in Indonesia in which very little research on this topic has been conducted (most of them were about creating or adopting a learning environment scale). Original research of this kind is relevant because school climate is a factor that enhances the effectiveness of a school in terms of student learning outcomes and progress (Anderson, 1982; Brookover et al., 1978; Creemers & Kyriakides, 2008; Fraser et al., 1988; Reynolds et al., 2015; Scheerens et al., 2003). School climate is also a malleable aspect of education that a school or local government can manipulate (Voight et al., 2013; Wang & Degol, 2016). So, arguably, it can be intervened immediately (although the effects may not occur immediately) without having to wait for a long time to change the policy. Thus, it is timely to examine school climate to inform new policy development and local government.

School climate is often considered similar to school culture and school ethos, such that some researchers used these terminologies interchangeably. However, school climate is unique in a sense that climate indicates a greater focus on how school members perceive the organisational and social climate of their school, whereas school culture looks for what members believe and think about the school (Van Houtte, 2005). Frequently, school climate is explored in terms of behavioural as well as perceptual terms, whereas culture is more about the values and norms held by school members (Hoy, 1990). Moreover, in distinguishing between the two terminologies, Hoy et al. (1991) postulated that school climate is a construct from a psychological standpoint, and school culture is a construct from an anthropological viewpoint. In contrast, the term school ethos seems to emphasise on how school members work together (Glover & Coleman, 2005). The present research adopts the point of view that school climate is the perception of school members about their school activities (i.e., teaching, learning process, or the relationship among school members). A detailed discussion about the

differences among these constructs is provided in the literature review chapter of this dissertation.

Educators have discussed and studied the school climate for more than 100 years (Anderson, 1982; Cohen et al., 2009). It is considered a crucial aspect of school effectiveness (Creemers & Reezigt, 1999; Mortimore, 1988; Scheerens, 1992; Scheerens & Bosker, 1996). It has a meaningful influence on student emotional development, self-esteem, self-concept, psychological well-being, and decreased student absenteeism (Thapa et al., 2003).

Scholl climate has been studied extensively in Western countries such as in the United States (US). However, only a few studies have been conducted in non-Western Countries such as Asian countries (Yang et al., 2013). For instance, in Malaysia, Adel and Zainal Ariffin (2011) examined the relationship between school climate and how teachers perceived the leadership style of the headmaster of their school. Faour (2012) investigated the relationship between school climate and citizenship skills in Arab countries. The study by Faour (2012) is related to the present study in some respect because he assessed the relationship between school climate and student non-cognitive skills. Specifically, Faour (2012) used a combined database of TIMSS 2007, PIRLS 2006, and PISA 2009. School climate research is also emerging in China. Yang et al. (2013) compared Chinese and US students' perception of their school climate. Jia et al. (2009) explored the relationship between school climate and student academic and emotional adjustment in China and the US. In brief, these studies in China revealed that compared to students in the U.S., Chinese students had a higher level of support from their teacher as well as their fellow students and that they had better opportunities to express their self-autonomy in the classroom. These two studies are particularly important in guiding the present study because it showed that students from Western and Eastern countries seem to emphasise a different aspect of school climate. However, how such findings also apply in the Indonesian context requires further clarification.

School climate research in Indonesia is scarce. For example, Wahyudi and colleagues developed the school climate instrument (Wahyudi & Darrell, 2006; Treagust, 2004), but their studies were conducted only from the point of view of teachers. Damanik & Aldridge (2017) assessed the relationship between school climate, headmaster leadership, and teacher self-efficacy. But this research too was conducted only from the point of view of the teachers. None of these studies focused on the relationship between school climate and the broader range of student learning outcomes. Up to date, no study has attempted to incorporate both student academic achievement and their social-emotional development. This study, therefore, seeks to explore the relationship between school climate and a broader range of student learning outcomes in Indonesia, where research on this topic is lacking.

1.3 Rationale for research

1.3.1 The importance of linking school climate with a range of learning outcomes as a means to promote school effectiveness

Studies on school effectiveness have mostly looked at the cognitive aspect of academic achievements, such as on mathematics or English (Kyriakides & Creemers, 2008; Opdenakker & Van Damme, 2005; Sammons et al., 1995; Slavin & Lake, 2008). Such focus on cognitive ability makes sense since the main objective of schooling is indeed to improve the way of thinking and habit of thought of the educational subject (Gardner, 2006). However, such emphasis means that little endeavour has been given to the non-cognitive outcomes of schooling. Indeed, up to date, there is no conventional standard measurement of the non-cognitive outcomes of schooling and valid and reliable instruments are scarce (Hattie, 2012; Luyten et al., 2005).

Such emphasis on cognitive outcomes has led some scholars to regard school effectiveness research as mechanistic (e.g., Elliott, 1996). Such a mechanistic approach would provide less assurance of schooling goals (Creemers & Kyriakides, 2010; Knuver & Brandsma,

1993; Mortimore, 1988; Thomas et al., 2000). Such a mechanistic approach is also ironic considering that student achievement is a multidimensional construct comprising both cognitive and non-cognitive aspects of schooling (Guskey, 2012). Indeed, the classic study by Mortimore and colleagues (1988) showed that 77% of the teachers from 50 schools in London had a social-affective purpose for their students. The examination of such affective domains is thus just as crucial as the cognitive domain of schooling (Reynolds et al., 2014).

According to Cohen (2006), the objectives of schooling are improving students' academic success as well as their social-emotional development. In the same way, Bloom (1976) mentioned that the teaching-learning process in schools is not for merely fostering cognitive development but should also promote students' affective skills. The outcome of affective aspects of learning is related to the way learners feel, their thoughts and behaviours, beliefs and all the processes they experience in the school on each subject. These affective learning outcomes can be assessed from the students' motivation, self-confidence, self-concept, and other attributes of the personal development of students (Guskey, 2012). These non-cognitive aspects of learning are not only important in their own right but are also intended to promote academic success (Baker et al., 2013; Elias, 2003; Watson et al., 2012; Zins & Elias, 2007).

UNESCO (2004) also highlighted that education should provide attention to both academic and affective (i.e., emotional, creative capacities, personal development) outcomes to allow students to reach their fullest potential. This strong view on developing academic and affective outcomes is consistent with Indonesia's education goals as required in the latest Education Legislation, Act Number 20 of 2003 on the National Education System National (Indonesia, 2003), which aims to develop both outcomes. Also, UNESCO (2014) emphasise the importance of non-academic learning outcomes as one of the Post-2015 Education Indicators to promote students' success. According to this view, it is essential to include both

student academic achievement and their affective outcomes in researching school climate as a particular key process related to a school's effectiveness. For that reason, this research consists of both aspects of education as dependent variables. The two learning outcomes are also viewed as interdependent as suggested by previous research (Marsh & Martin, 2011; Marsh & O'Mara, 2008; Parker et al., 2014; Paschal, 1968; Seaton et al., 2014).

Indeed, schools are different in promoting student success, both academic and affective outcomes as Reid et al. (1987, p. 4) pointed out:

> "*Two students from similar backgrounds and similar intellectual abilities can perform differently at two outwardly similar schools because of the unique blend of academic and social circumstances to be found within the two establishments.*"

This statement alludes to the school climate. It can enhance student outcomes (Creemers & Reezigt, 1999; Mortimore, 1988; Sammons, 1999; Sammons, 1999; Scheerens, 1992). Indeed, effective and ineffective schools were found to have different climates' (Creemers & Reezigt, 1999, p. 36).

School climate is also one of the key process criteria used to evaluate school effectiveness in the early birth of the SER field (Brault et al., 2014). For example, Brookover et al. (1978) emphasised the importance of school climate in promoting student learning. Accordingly, Edmonds' (1979) early model of effective schools recognised the important role of school climate. He suggested that powerful school leadership, high expectations of academic performance, safe and well-ordered environments, the emphasis on basic academic skills, and a pupil progress surveillance plan represent key processes that underpin a school climate that encourages academic success. Hoy et al. (1991) have also acknowledged that the overall school effectiveness is affected by a highly positive and favourable school climate.

Although school climate has been identified as essential for enhancing student learning outcomes, only a few studies of school climate have been conducted within the Asian context, particularly in Indonesia. More research in this field is needed in the Indonesian educational context. School climate in Indonesia may have specific differences from Western countries in which school climate is often researched.

Indonesian people are typically collectivist and maintain traditional community rules and values; group concerns are more central than personal preferences and interests (Hofstede, 1993; Hofstede & Minkov, 2010). For example, Batak Toba (one of the major ethnic groups in North Sumatera) sees a higher value in education than some other ethnic groups (Irmawati, 2007). Parents will do their best to send their children to school rather than fulfilling their need even when they have very limited money. Considering the unique nature of the Indonesian culture, it is expected that these features will be reflected in the profile of the school climate and the relationship between school climate and student learning outcomes. The Batak Toba people represent a single culture in Indonesia, while Indonesia is diverse in cultures, ethnics, languages, religions and local authorities (Novera, 2004). This suggests that within-country cultural variation in Indonesia may also influence school climate.

The present study seeks to address a gap in the literature which requires exploration and clarification about Western features of school climate in the Indonesian context. The present study also investigates the links between school climate and a range of educational outcomes that is a central concern of school effectiveness research. It contributes to the school effectiveness body of knowledge in analysing a particular aspect of the school process - school climate - in a developing society such as Indonesia. By doing this research, countries which have similar characteristics may learn from the results obtained in Indonesia.

This research also responds to other criticisms on SER, which has been argued as having lack of theory-driven and methodologically criticised as too “mechanistic” and naively

"positivistic" (Angus, 1993, p. 335). Therefore, the present research utilises a newly developed SER theoretical framework (Dynamic Model of Educational Effectiveness (DMEE) – which will be explained later in Chapter 3) as guidance. In which quantitative and qualitative methods are utilised.

In the next section, the author will advance to discuss the rationale of conducting the present research in the Indonesian context. The discussion will begin with the education regulation in Indonesia and its relation to the influence of a positive school climate in supporting better learning outcomes.

1.3.2 Indonesia Education Act and creating a positive school climate: The national context rationale

The Indonesian Regulation on the National Education System (Indonesia, 2003) aims to develop children's fullest potential. This law is important to establish a constructive learning environment. This regulation aims to incorporate students' and parents' goals at the forefront of education. Such an approach would motivate school members to be more enthusiastic to give their best. It also generates a sense that they are being valued and emotionally connected to the school. Freiberg and Stein (1999) labelled such a condition as school climate. The Indonesian Regulation on the National Education System has been implemented almost for two decades, and as mentioned previously, there has been no systematic effort to evaluate its implementation. This makes it imperative to conduct school climate research in Indonesia.

In a recent empirical review, Thapa et al. (2013) emphasised that school climate: (a) has a powerful effect on students' learning motivation (Eccles et al., 1993a); (b) reduce the negative influence of socio-economic background on students' academic attainment (Astor et al., 2009), (c) performs as a 'protecting factor' for young people's self-development and learning (d) contributes to fewer harassments at school (Eliot et al., 2010; Reaves et al., 2018).

The examination of these factors in Indonesia is important for several reasons described in the following.

First, reports from TIMSS and PISA have consistently ranked Indonesia as one of the lowest in terms of international achievement (Mullis et al., 2012; OECD, 2013). This makes it imperative to identify approaches to improve the educational standard in Indonesia. Such identification can be made by analysing TIMSS 2011 database and by combining it with the stakeholders' views.

Second, most Indonesian students live in an economically disadvantaged family. They have lower access to obtain better education (i.e., qualified teachers, laboratory equipment, textbooks). Kaluge et al. (2004) found that there is a significant discrepancy between schools in Indonesia. Some schools are outstanding, but the majority are poorly managed. Such a difference is partly because many students are from low-income families (World Bank, 2014b). Therefore, researching school climate may be beneficial in identifying evidence to inform possible new school improvement strategies for Indonesian schools that seek to decrease the negative impact of these economic disadvantages.

Third, the examination of school climate has the potential to assist the formation and development of a positive learning environment. A positive learning environment can act as a protective factor for students' learning. It encourages students to go to further education.

Fourth, school harassment is problematic in Indonesia. Former Minister of Education and Culture of Indonesia, Baswedan (2014), pointed out that school harassment such as physical, emotional, or sexual abuse often come up in the news across the country. The data from the International Centre for Research on Women (ICRW) backed this claim (Bhatla et al., 2014), making it even more imperative to study school climate factors that might reduce such undesirable outcomes at school.

Fifth, Indonesia has two responsible bodies for supervising and managing the education system (Indonesia, 2003). The Ministry of Education and Culture (MOEC) manages approximately 81% of general schools, while the Ministry of Religious Affairs (MORA) supervises 19% of Islamic schools or *madrasah* (Ministry of Education and Culture, 2017a). These two ministries run two different curricula and are very different in resources (e.g., teachers' qualification, school infrastructure, etc.). The *madrasah* typically has lower resources than the general school (Ali et al., 2011). Considering the lack of systematic examination on how these two types of schools differ in terms of school climate, it becomes imperative to compare the two schools. This research may inform a novel approach to improve the quality of education in Indonesia across the two ministries.

Finally, Indonesia has also a decentralised education system, in which the local government has a diverse capacity in managing and delivery of education services (Al-Samarrai, 2013; Ministry of Education and Culture, 2013). This system made a variety of gaps at the province and district level (e.g., teacher training, resources, and finance: Yeom et al., 2002). The present research would help local policy development to improve the quality of education of both the general school and *madrasah*.

This research is original and significant considering that there has been no research that investigates the relationships among school climate, learning outcomes, student achievement, and self-beliefs in Indonesia. This research also employs a mixed method, combining quantitative and qualitative methods to develop more substantial explanations of how and why variables are interrelated. Previous research has not implemented such an integration of methods in Indonesia. Specifically, the author examines TIMSS 2011 data and combines it with case studies to provide a more comprehensive description of school climates and how it influences students' outcomes.

As explained before, school climate is the 'personality' of a school that has a similar characteristic of a living organism (Freiberg & Stein, 1999; Halpin & Croft, 1963). Considering that every school operates as an autonomous organisation, it will be meaningful to find the pattern to distinguish different school climate and its influence on student learning. The present research would also provide novel evidence to support policymakers and educators in understanding the importance of a positive school climate to enhance student achievement and their psychological development in Indonesia and similar developing countries (e.g., Thailand, Brazil, etc.).

1.3.3 Personal rationale

As an educational psychologist, my professional work has typically involved examining the impact of schooling on students' social-emotional development. My previous research, for example, explores the association of classroom climate on student achievement (Tarmidi & Wulandari, 2005). I have explored the application of cognitive behaviour therapy to enhance student's academic self-concept (Tarmidi & Akbar-Hawadi, 2009), the relationship between self-regulated learning skill and parents' social support (Tarmidi & Rambe, 2010), and the relationship between emotional intelligence and academic resilience (Dazeva & Tarmidi, 2013; Pulungan & Tarmidi, 2012; Tarmidi & Vanita, 2008). My interests to improve students' achievements underline my enthusiasm to study the academic and affective outcomes of education.

I am also a lecturer at a university, specialising in educational psychology. As a lecturer, I must engage in "*tri-dharma perguruan tinggi*", loosely translated as the three primary duties of a lecturer. Besides teaching and research, I also need to take part in community development. I have contributed to assessing the quality of education at schools in the province of North Sumatra. The present research offers me a chance to better understand and learn a novel

approach to evaluate school quality. I hope to improve school effectiveness in the province of North Sumatra by encouraging a better school climate.

1.4 Research aims and objectives

The present study aims to examine the relationship between school climate and a range of student outcomes. Including academic achievement and students' self-beliefs in Indonesia. The study builds on the literature of international school effectiveness and school climate to explore the relationship between school climate and student outcome variables taken from TIMSS data of Indonesia. The study also seeks to examine school climate differences across different schools (general and *madrasah*) and their effect on student learning outcomes as perceived by school stakeholders. This research also examines whether school climate can explain how a school can be an effective school, refers to schools that can promote a broader range of educational outcomes, including academic and non-academic (Sammons, 1999). This research viewed that the effectiveness of a school is an integrated element in the school system. It aims to 'identify the interactions between key factors that operate at the school, classroom, and individual student levels, and their contributions to student performance' (Creemers & Kyriakides, 2008, p. 4).

This research uses data from TIMSS 2011 to measure student academic achievement in mathematics (Mullis et al., 2012). The author also assesses students' self-beliefs by measuring the students' self-concept and self-efficacy in mathematics. The assessment of students' beliefs would also be obtained from the TIMSS 2011. Self-concept and self-efficacy will be used because they are powerful predictors of long-term academic achievement (Parker et al., 2014). Long-term academic achievement is essential to keep students in continuing their formal schooling, as students who achieve less are more likely to drop out (UNESCO, 2014).

Many Indonesian students come from disadvantaged families (World Bank, 2014b). They are at risk of school dropout. Although Indonesia has rapidly expanded access to junior

secondary schools, about 1.74% or 212,085 students in lower secondary school dropped out, and 7.95% of graduate students or 79,209 did not continue their education to the next level (UNESCO, 2014; Al-Samarai, 2013; World Bank, 2014a; Bjork, 2004; Ghozali, et al., 2013; Saputro et al., 2018). Students from low-income families dropped out of school because their parents do not have financial means to support their children for schooling (Mo et al., 2013; Makwinja-Morara, 2009; Traag & van der Velden, 2011). Whether students' self-belief could nullify or reduce the effects of economic disadvantages on school dropout is examined in the present research.

To clarify, the objectives of the present research are as follow:

1. To conduct a literature review to explore and critique relevant literature in conceptualising school effectiveness, school and classroom climate and its effect on students' learning outcomes, particularly students' self-beliefs and academic performance.
2. To describe the key and relevant features of the Indonesian research context.
3. To conduct a multilevel statistical analysis of secondary data from Indonesia from TIMSS 2011 related to maths achievement, school climate and students' self-beliefs.
4. To conduct a qualitative case study analysis of four different schools in term of stakeholders' perceptions of the school climate, effectiveness and context, and to compare in particular the association between school types (general and Islamic), school climate and student's outcomes concerning mathematics achievement and academic self-beliefs.
5. To draw on the research findings to provide recommendations that might assist Indonesian and other similar developing country schools in developing new strategies aiming to improve education quality by promoting school climate and students' self-beliefs.

6. To contribute to the knowledge base in school effectiveness, specifically in presenting recent evidence on school climate and its relationship to students' achievement and their self-beliefs in a newly developing country context - Indonesia.

1.5 Research questions

The researcher utilises the following specific research questions (RQ):

RQ1: What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?

- RQ1.1: What is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?
- RQ1.2: After controlling for student, teacher, and school characteristics, what is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?
- RQ1.3: What are the school climate factors that significantly explain the variance between school and classroom performance among Indonesian Year 8 students in math and self-beliefs before and after adjusting the characteristics of the student, teacher, and school?
- RQ1.4: Why do students from the general school have higher achievement and self-beliefs than those from *madrasah* before and after controlling school climate and other factors?

RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)?

- RQ2.1: (a) What are the similarities and differences in school climate between high and low-performing schools? (b) and between religious/non-religious school settings?

- RQ2.2: Are there new factors that can be obtained from a qualitative inquiry that are relevant to highlight differences between the high and low performing schools?

1.6 Overview of the theoretical framework

This study seeks to examine key elements of school climate and its relationship with student academic and affective outcomes. The study draws upon the concepts, theories, and paradigms developed within the field of school effectiveness. For example, the dynamic model of educational effectiveness applies to the current study because the model incorporates school level aspects related to learning outcomes (Creemers & Kyriakides, 2010a). The model is also one of the most dominant and updated theoretical concepts in the field compared to the previous models (Scheerens, 2013; Teddlie & Reynolds, 2000).

There are some reasons to draw on key elements of this model. First, the dynamic model considers a broader range of educational outcomes rather than being limited to academic achievement (Creemers & Kyriakides, 2006). It is in line with this research that seeks to explore academic and affective outcomes of education. Second, this model considers five evaluative dimensions (frequency, focus, stage, quality, and differentiation) concerning the various school processes that contribute to school climate and effectiveness. By considering these five dimensions, it is arguably possible to explain in more depth why some schools appear to be more effective in promoting student learning outcomes than others, since key relationships may not be necessarily linear (Creemers & Kyriakides, 2006).

The *frequency* dimension refers to the number of activities associated with the school effectiveness factors, while the other four dimensions measure the quality of each of the school effectiveness factors. The *focus* dimension deals with whether the school effectiveness factors are too narrow or too broad. The *stage* dimension refers to how long the school effectiveness factors have been applied. *Quality* refers to activities or policies that have been applied to ensure that the school effectiveness factors are functioning as they should be. Finally,

differentiation concerns the degree to which school effectiveness factor-related activities are performed consistently in the same manner (Creemers & Kyriakides, 2008). The detail of the operational descriptions of the five evaluative dimensions is described in Chapter 3.

This study chose a framework from a recent empirical review by Thapa et al. (2013). They summarised five key factors that contribute to the overall school climate construct. The five factors include (1) *safety factor*, which refers to social-emotional and physical safety in school, (2) *teaching-learning* factor, which refers to teaching practices within the school and classroom and outcomes expectations of teachers and students, (3) *relationship factor*, which highlights the social interactions between school members, (4) *physical environment*, which relates to physical resources of learning, and (5) *school improvement processes*, which focuses on how school principals improve school processes (Cohen et al., 2009; Thapa et al., 2013).

The review is based on and extends the previous school climate review by Cohen et al. (2009). This theoretical framework is based on the most extensive empirical review on school climate. It is also comprehensive in the sense that it includes almost all of school climate factors referred to in past research. As stated before, the factors of school climate included in this review are agreed by most researchers (i.e., Cohen et al., 2009; Thapa et al., 2013; Zullig et al., 2010; Wang & Degol, 2015). The details of the chosen framework can be found in Chapter 3.

1.7 Overview of the methodological framework

To address the research questions and to achieve the stated aim and objectives, the study integrates quantitative and qualitative methods to present a more comprehensive portrayal of a school climate and the links to the student outcomes in different school types (Teddlie & Sammons, 2010; Teddlie & Tashakkori, 2008). Therefore, a mixed-methods design is applied. This approach is robust in the sense that quantitative measurements of school climate alone are limited because they only offer numerical data to represent the complexities of school functioning. Qualitative methods, specifically case study research comprising interviews and

focus group interviews will provide a more contextualised and evidentiary description of school climate and student learning besides the numerical results. Integrating those multiple types of data may provide a better understanding of the research problem (Tashakkori & Teddlie, 2003). In studying SER, Scheerens and Bosker (1996) recommended combining large-scale datasets analysis (including TIMSS, PISA, PIRLS) with more in-depth data collection (e.g., qualitative data). Because the mixture of both robust statistical findings as in TIMSS study and in-depth descriptions of specific cases showing those findings has the potential to deliver new insights and strengthen understanding of EER topics that neither can achieve alone (Sammons, 2010).

The methodological framework, together with a discussion about assumptions and the choice of a pragmatic philosophical paradigm, mixed-method research design, methods, and data analysis justification, is explained further in Chapter 3.

1.8 Brief description of the thesis by chapters

The author organises the rest of this thesis into six chapters. Chapter two places the research in the context: the Indonesian education system. It starts by reviewing the educational policy to get a general idea of how schools in Indonesia operate. Followed by examining the existing strengths and weaknesses of the Indonesian education system critically.

Chapter three presents the theoretical background of the study, by reviewing the key concepts and previous research in three areas relevant to the aims of this study: (1) key concepts of SER, particularly dynamic model of educational effectiveness (DMEE) and (2) key concept of school climate. The author ends Chapter three by outlining the analytical tools of this study derived from the literature and by enquiring the generalities of these bodies of knowledge to other country contexts, in particular, the Indonesian context.

Chapter four presents the method, demonstrating how and why the research approach and methods have been selected, giving particular attention to the philosophical,

epistemological view of pragmatism which underpins this study. This chapter also explains the use of DMEE and Thapa et al. (2013) and other previous research in developing the research instruments for data gathering and data analysis framework and ends by reviewing the ethical issues and methodological limitations of the study. Chapter five presents and discuss the quantitative results, to answer RQ1 (oriented to assess the range and extent of the relationship between school climate and academic performance in Indonesian Year 8 students using TIMSS 2011 data).

Chapter six describes the qualitative results, addressing RQ2, contrasting the perspectives and opinions of school members (headteacher, teachers, and students) in four illustrative case schools which vary in terms of academic performance and general/religious type. The findings examine school climate and how school climate influences student learning outcomes. Special attention is given to the reporting of the five effectiveness evaluative dimensions of DMEE in analysing school climate factors. These evaluative dimensions are used to describe and differentiate school practices in the five school climate factors. The author also addresses the identification of new school climate factors that emerged from the data analysis. Chapter seven (discussion and conclusion) outlines and reviews the key quantitative and qualitative findings concerning previous research, limitations, and discusses the potential for new research in this area.

Chapter 2: The Indonesian context

2.1 Introduction to the chapter

In this research, Indonesia is the key point of interest. This chapter puts the research problem within the Indonesian context. The first section of this chapter concisely introduces Indonesia's geographical location, population, and culture. Subsequently, Indonesia's education system and the challenges in implementing its education system are presented. This chapter also explains the problems that prevents the *madrasahs* from achieving the same level of effectiveness as achieved by the general schools.

2.2 Indonesia geographical location and culture

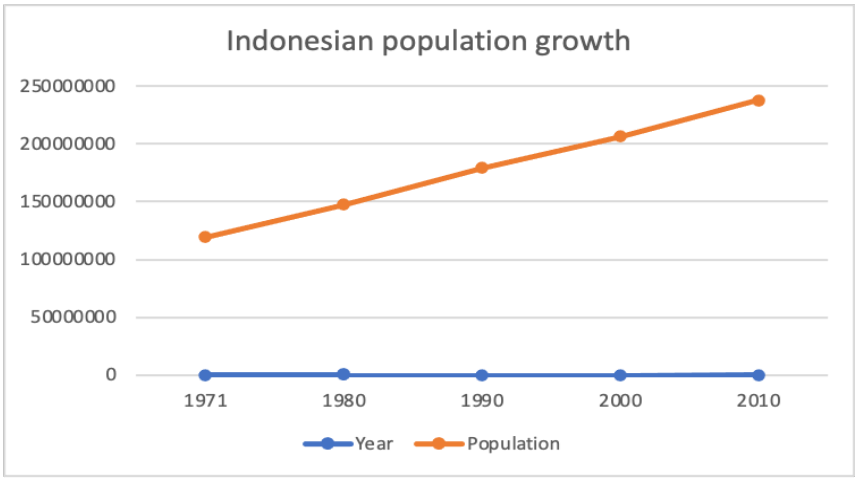
2.2.1 Geographic location and people

Indonesia is positioned on the equator, located between mainland Asia and Australia (see Figure 2.1). It is one of the biggest archipelago countries (comprising 18,000 islands) in the world, which covers the length of one-eighth of the earth's circumference across the equator (Legge et al., 2016).

Source: maps.google.com

Figure 2-1. Map of Indonesia

Indonesia has the world's fourth-largest population after China, India, and the United States. According to the Indonesian Central Bureau of Statistics or BPS (2012), the population of Indonesia was about 237 million in 2010, almost double compared to 1971 (see Figure 2.2). It is projected that by 2045, the total population would increase to 318.9 million people (BPS, 2018).



Source: BPS (2012)

Figure 2-2. Indonesian population growth

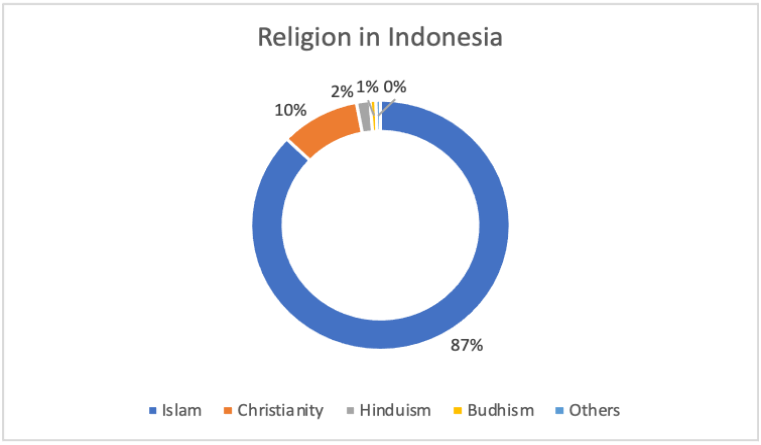


Figure 2-3. Religion in Indonesia

Approximately 87 per cent of Indonesians are Muslims (see Figure 2.3), making Indonesia the world's largest Muslim country in the world. The country is a diverse, complex, multi-ethnic, and economically heterogeneous (Hugo, 1995). The people are varied, from rural

hunter-gatherers to a modern urban elite. In terms of language, it has an outstanding ethnic diversity with over 300 local languages (World Bank, 2018).

According to the Indonesian Central Bureau of Statistics (2012), there are three largest ethnic groups in Indonesia. The Javanese make up about 40% of the population. They primarily live in Java, the world's most inhabited island, and home to over 50% of the overall Indonesian population. The Sundanese make up about 15% of the population. They primarily live in Western Java. The Malays make up about 4% of the population; they spread across Sumatera, Kalimantan, and Sulawesi. Jakarta as the capital city is inhabited by 10 million people.

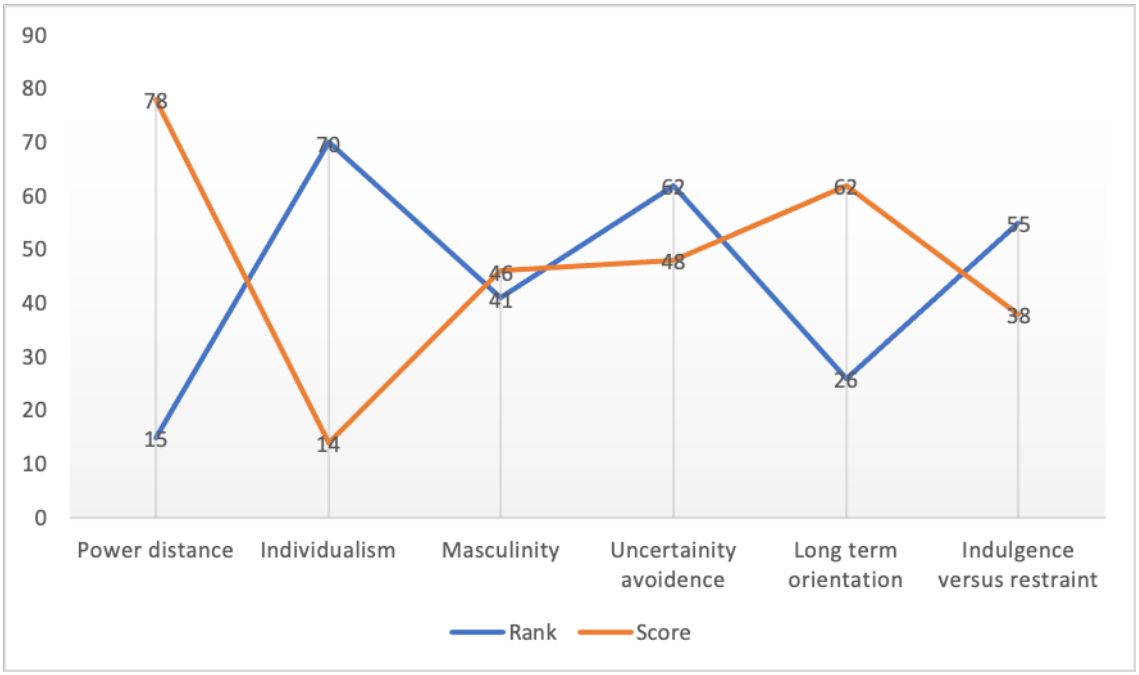
2.2.2 The Indonesian culture

Indonesia is mixed in term of religions, ethnicities, customs, languages, and races (Wihardit, 2017). Studies suggest Indonesia as a collective society (Magnis-Suseno, 1997; Oerter et al., 1996). For instance, the Javanese people hold the concept of *rukun* as their principle to maintain social harmony and communality (Rahmi et al., 2001). *Rukun* is a practice to orient mutual understanding, respect, and social adaptation among individual members of the Javanese society. Although each sub-culture in Indonesia has their unique term, maintaining harmony is a collective value shared by most sub-culture in Indonesia (French et al., 2005; Forshee, 2006; Magnis-Suseno, 1997; Noesjirwan, 1978).

While cultural conception such as harmony maintenance is a cultural strength, some cultural conceptions may hinder interactions in the educational setting (Dardjowidjojo, 2001). For instance, students would not dare to challenge their teacher. A teacher's word is like an order. Challenging a teacher's view is disrespectful and may cause disharmonious relations.

The research by Hofstede and colleagues are useful to describe Indonesian national culture in a general sense. Hofstede's explained Indonesian culture by looking at six cultural dimensions (i.e., power-distance, individualism, masculinity, uncertainty-avoidance, long-term orientation, and indulgence (Hofstede, 1986; Hofstede et al., 2010). Figure 2-4 shows the

Indonesian cultural index ranks relative to other 76 countries. Indonesia has a high power-distance but low in individualism and appeared to be somewhere in the middle in terms of the other four cultural dimensions. In the next section, the author shall briefly discuss each of the cultural indicators.



Source: compilations from Hofstede et al. (2010)

Figure 2-4. Indonesian culture index

2.2.2.1 Power distance

Power-distance describes how low-status members of a society accept and expect inequality of power in their society. (Hofstede, 1984; Hofstede, 1986; Hofstede et al., 2010). In a low power-distance society (e.g., Western societies) people value equal opportunities and rights, and power is decentralised (Mooij, 2014). It is normal for leaders to consult and involve their subordinates before deciding. But the interactions and communications are mostly participative and direct. In the educational context, teachers respect students' independence and teaching process is usually student oriented. The students are also allowed to challenge their teachers' point of view (Hofstede, 1986).

Conversely, in a high-power-distance society, the rights and privileges between the power holders and their subordinate are unequal (Hofstede, 1984, 1986). The distribution of power is often centralised and those with power are mostly directive. They expect their subordinate to follow their will.

In Indonesia, the educational process is usually directed by the teacher (teacher-centred). Students expect their teachers to start an interaction in the classroom. Students participation depends on the teacher's order. Teachers are rarely contradicted or critiqued (Hofstede, 1986). In most secondary school classrooms, teachers predominantly lead the conversation, the student must listen and obey their teacher (Buchori, 2001; Sopantini, 2014). Questioning is observed as to challenge a teacher's authority, and a show of one's ignorance (Lewis, 1997).

2.2.2.2 Individualism

Hofstede categorised Indonesia as a collectivist society. Such society highly endorses group harmony and concern for others (Hui, 1988). A person in a collectivist culture is expected to be closely knitted and integrated into a strong and cohesive group, primarily with the person's nuclear family, extended family, and the community (Hofstede, 1986). Individuals are expected to put aside their self-interest and follow the views of the group that he or she belongs to. Like most Asian cultures, the cultural value in Indonesia emphasised interdependence (Geertz, 1973). In short, Indonesians are people with prime interest in the community rather than individuals (House et al., 2013).

In a more recent classification of cultures (Globe 2020; globeproject.com), Indonesia is grouped alongside India, Iran, Malaysia, Philippines, and Thailand as a collectivist society (high collectivism). Being high in collectivism also influences school processes in Indonesia. A harmonious learning environment should always be maintained. Student speaks only when

the teacher requests them to speak (Hofstede, 1986). Most teaching-learning practices in schools or classrooms are teacher-centred.

2.2.2.3 Masculinity

Compared to other countries in East Asia, Indonesia has a low masculinity index. But compared to the general world population, Indonesian have a medium to high masculinity score. High masculinity index implies that society is much more determined by material successes and competition (Hofstede, 1984). Low masculinity implies that society values life quality above achievement and focuses primarily on interpersonal relationships (Hofstede, 1984). In the school context, Hofstede (1984) identified that in a high masculinity society, teachers openly praised reward students who have high academic performance. The teacher usually makes such students a role model for other students. Students who are not excelling in academic achievement often feel uncomfortable because the school or teacher does not appreciate alternative talents (e.g., athleticism or musicality).

The greater focus on student academic achievement is in line with the research by Revina (2017). She examined the influence of culture on maths education in Indonesia. She observed that teachers in the math classroom focus far more on the right answer than on helping the student understand the task. In another study that aimed to compare learning patterns between Indonesian, Sri Lankan, and students from The Netherland, Marambe et al., (2012) reported that memorisation and rehearsing techniques were used mostly in Indonesia compared to students in Sri Lanka and The Netherland. The learning orientation of Indonesian students is also more certificate oriented.

2.2.2.4 Uncertainty avoidance

A high uncertainty-avoidance index is associated with a high-stress level, the needs of punctuality and precision for encountering of an uncertain future (Hofstede, 1984, 1993). Indonesia has a moderate uncertainty avoidance index. Meaning that Indonesian are moderate

in terms of general stress level, the need to be punctual and precisions. This may be because of the principle of peace or harmony (*rukun*) mentioned in the previous, in which people avoid conflict or confrontation (Guinness, 1986; Magnis-Suseno, 1997).

Societies with low uncertainty-avoidance index are usually unstructured, as they have no rigid schedules. They also have a wide range of activities. The teacher often uses direct and straightforward language in the classroom. In contrast, in high uncertainty-avoidance societies, the learning environment is predominantly organised, with realistic goals, strict timetables and comprehensive tasks. The teacher uses academic language from time to time, and the educator considers themselves as experts. Indonesia is somewhere in between these two distinct cultures (Hofstede, 1986).

2.2.2.5 Long term orientation

Indonesia's score on this measure is relatively high. Meaning that they have a pragmatic approach to achievement. The long-term emphasis means that they can adapt to change rather quickly and maintain perseverance to achieve results (Mooij, 2014). In short, Indonesian people are likely to have a good orientation about their future.

2.2.2.6 Indulgence

Indulgence describes how a society is enjoying life (Hofstede, 2011). Concerning the indulgence aspect, Indonesia has a relatively low score in this index. It means that as a controlled society, Indonesian people viewed social norms as a restriction. They often feel hesitant to be indulgent.

To sum up, the educational system in Indonesia may be reflected by its culture as a nation. In most of the secondary school classroom, teachers are dominant, and most of the classroom instruction is teacher-centred. This collectivist and high-power index society also expect students to obey their teachers as authoritative figures. Students are also trained in “how to do” and taught that there is only one acceptable viewpoint to answer a question. Both

students and teacher insist to treat the teachers in respectful ways. As such, the students avoid debates, discussions, and arguments in the classroom and only follow the instructions given by their teacher (Maulana et al., 2016).

The work of Hofstede and colleagues are interesting and intuitive (Hofstede, 1984, 1986, 2011; Hofstede et al., 2010), but their work has also been criticised (Baskerville, 2003; McSweeney, 2002; Signorini et al., 2009). For instance, McSweeney (2002) asks a fundamental question, "Do nations have cultures?" (p. 89). Another critique argued that Hofstede neglected the complexity of culture (Signorini et al., 2009). As such, the description of Indonesian culture based on Hofstede's work in this study is aimed to get a brief understanding only. Not intended to describe an entire Indonesian culture. Moreover, the categorisation was made based on employees of the International Business Machines company (IBM) and was conducted almost 50 years ago (Signorini et al., 2009). Considering the flexibility and fluidity of culture (Forshee, 2006), the current Indonesian culture may have changed.

2.3 Indonesia National Education System in brief

The World Bank (2014c) reported that the Indonesian educational system is the fourth largest in the world after China, India, and the United States. The nation has over 250,000 schools, 50 million students, and over 2.6 million teachers that spread over in 34 provinces and 514 districts (Ministry of Education and Culture, 2017a). These provinces and regions are geographically dispersed into five main islands: Sumatera, Java, Kalimantan, Sulawesi, and Papua.

The schooling system in Indonesia is managed by two ministries, the Ministry of Education and Culture (MOEC) and the Ministry of Religious Affairs (MORA). MOEC primarily oversees public and private general schools, while MORA supervises public and private Islamic school (Ministry of Education and Culture, 2017a). There are 81% of primary

and secondary schools under the MOEC and the remaining 19% are under MORA (Ministry of Education and Culture, 2017a, see also Figure 2.5.). This unique dual structure exists since colonial times. At that time, public schools were targeted at educating the Dutch elite, while Islamic schools pursued to teach the broader population (Ali et al., 2011).

The Indonesian system of education comprises three core levels: basic, secondary, and higher education (Ministry of Education and Culture, 2012). Basic education comprises six-years of primary school (Year 1-6), followed by three years of lower secondary school (year 7-9). Secondary education covers a three-year schooling program (Year 10-12). There are two forms of secondary education general and vocational. Table 2-1 shows the Indonesian basic and formal education system. Higher education includes the undergraduate and postgraduate program. In the next section, the author shall discuss the three-level of the education system; basic, secondary, and higher education.

Table 2-1: Basic and secondary schooling system in Indonesia

Stage	Age	School Year	Level	Pathway	
				General School	Islamic School
Basic Education	7–12	Six years (Year 1-6)	Primary school	General primary school (*Sekolah Dasar-SD)*	Islamic primary school *(Madrasah Ibtidaiyah-MI)*
	13–15	Three years (Year 7-9)	Lower secondary school	General lower secondary school *(Sekolah Menengah Pertama-SMP)*	Islamic lower secondary school (*Madrasah Tsanawiyah-MTs*)
Secondary Education	16-18	Three years (Year 10-12)	Upper secondary school	General upper secondary school (*Sekolah Menengah Atas-SMA)*	Islamic upper secondary school *(Madrasah Aliyah-MA)*
				Vocational upper secondary school *(Sekolah Menengah Kejuruan-SMK)*	Islamic Vocational upper secondary school *(Madrasah Aliyah Kejuruan-MAK)*

Source: Compilation from Government Regulation Number 17 the Year 2010 of Educational Management (Indonesia, 2010)

2.3.1 Basic education

According to Government Regulation of Republic Indonesia Number 17 the Year 2010, about educational management, basic education is the basis of secondary education. The official minimum age to enter a basic school is seven years old but many 5 to 6-years old children already enrolled as a first-year student in primary school. This is especially common commonly in private primary schools (Barakat & Bengtsson, 2018). Student can choose between two different pathways of basic education, general primary schools or *Sekolah Dasar* (SD) and *Madrasah Ibtidaiyah (MI*) for Islamic primary schools. At lower secondary schools, a student can enrol in *Sekolah Menengah Pertama (SMP)* for general school or *Madrasah Tsanawiyah (MTs*) for those who wish to attend an Islamic school (Indonesia, 2003, 2010).

For the general education pathway, from Year 1 to Year 2, the curriculum contains civics education, religious and moral education, Bahasa Indonesia (Indonesian language), mathematics, art, and physical education. Starting from Year 4, sciences and social studies are added in the curriculum (Ministry of Education and Culture, 2013b). Those curriculums apply to both general and Islamic schools. The difference between the two types of school is regarding the religious and moral education. For those who attend the Islamic primary school, religious and moral education is emphasised. Specifically, starting at Year 4, they additionally learn Quran and Hadith, Islamic theology (*aqidah*), Islamic jurisprudence (*fiqh*), Arabic language, and Islamic history (Ministry of Religious Affairs, 2014a). As such, Islamic primary school has about 4-6 hours longer study time in a week compared to the general school.

The duration of secondary school is three years (Year 7-9). The curriculum is essentially the same as primary school, but with English and information technology added as additional teaching subjects (Ministry of Education and Culture, 2018). Students enrolled in *Madrasah Tsanawiyah* (Islamic lower secondary school), have 8 hours longer learning time in

a week because they have additional Islamic subjects (Indonesia Ministry of Religious Affairs, 2014a, 2014b).

2.3.2 Secondary education

Secondary education extends basic education (Indonesia, 2010). The duration is three years (Year 10-12). Secondary education has four pathways: The general upper secondary schools or *Sekolah Menengah Atas* (SMA), Islamic school or *Madrasah Aliyah* (MA), general vocational secondary schools called *Sekolah Menengah Kejuruan* (SMK), and Islamic vocational school or *Madrasah Aliyah Kejuruan* (Nor & Malim, 2014). Starting from the second year, students focus on one of four core disciplines. They can either focus on sciences, social sciences, languages, or religion (at Islamic school). A student selects the core discipline group based on their preferences. Those interests are determined by their preferred field to study at a higher educational level.

2.3.3 Higher education

Higher education comprises diploma or college, undergraduate, postgraduate, specialised postgraduate, and doctoral programs. The higher education takes form as an academy, polytechnic, institute, and university. These institutions could run academic, professional, vocational, or technical education programmes.

2.4 Challenges in Indonesia education system: existing strengths and weaknesses

Indonesia has a firm commitment to reform its education system (Wales et al., 2016). In term of net enrolment ratio (NER), which shows the numbers of students enrolled in primary school, Indonesia has achieved over 90% (UNESCO, 2004). Moreover, for the secondary school level, Indonesia has achieved 79% (World Bank, 2019). This means that the prominent educational challenge in Indonesia nowadays is not only to improve access (i.e., school enrolment) but also to improve the quality of education (Rosser, 2018; Wales et al., 2016). To

reach this end, the Indonesian government aims to establish a very ambitious goal, a world-class education system by 2025 (Rosser, 2018).

However, Rosser (2018) added that Indonesia has been unsuccessful in promoting high-quality education system. He claims this is not only because of poor management practices, inadequate finance support, and lack of labour resources but also about political will. For instance, the curriculum is usually altered each time a new government is formed (Bjork, 2004; Chang et al., 2014; Yeom et al., 2002). Such a rapid and often changing educational system is probably a reason for Indonesian low learning outcomes, as discussed next.

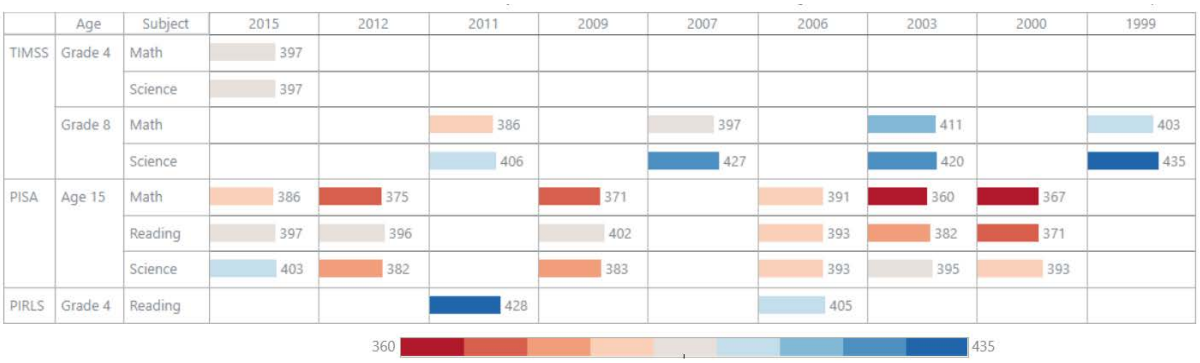
2.4.1 Low learning outcomes compared to other countries

The literacy ratings in Indonesia have been significantly improved. Efforts have also been made to reform the educational systems, in which significant improvements had been demanded in the quality of Indonesian education (Rosser, 2018; UNESCO, 2004; Wales et al., 2016). However, the condition becomes problematic since no official definition exists about how to measure the improvement (Tobias et al., 2014).

Tobias et al. (2014) informed that Indonesian policymakers mostly agreed with the use of international assessment, such as PISA and TIMSS test scores, particularly because of the increased criticism (i.e., massive cheating, the exams heavily focus on cognitive ability) of the national educational test system in recent years (Berkhout et al., 2019; Ministry of Education and Culture, 2013a). Although there are critiques on the use of these measurements (explain later in this Chapter), they are helpful considering that no better alternative is available. In the next paragraph, the author shall explain the quality of education in Indonesia utilising international measures.

By looking at the OECD (2013) report, Tobias et al. (2014) pointed out that Indonesia's PISA performance has been improved between the year 2000–2009. But the differences between the best and worst performers are also widening. Patterns in mathematics and science

achievement have become more ambiguous. Science achievement relatively decreased from 2006 to 2012. The findings of TIMSS also support a decline in science achievement during that time, with a significant drop in average test results during 2007–2011. Figure 2-6 shows the achievement progress of Indonesian students in different international large-scale assessment in education from 1999 to 2015.



Source: http://datatopics.worldbank.org/education/wDashboard/dqlearningcnty

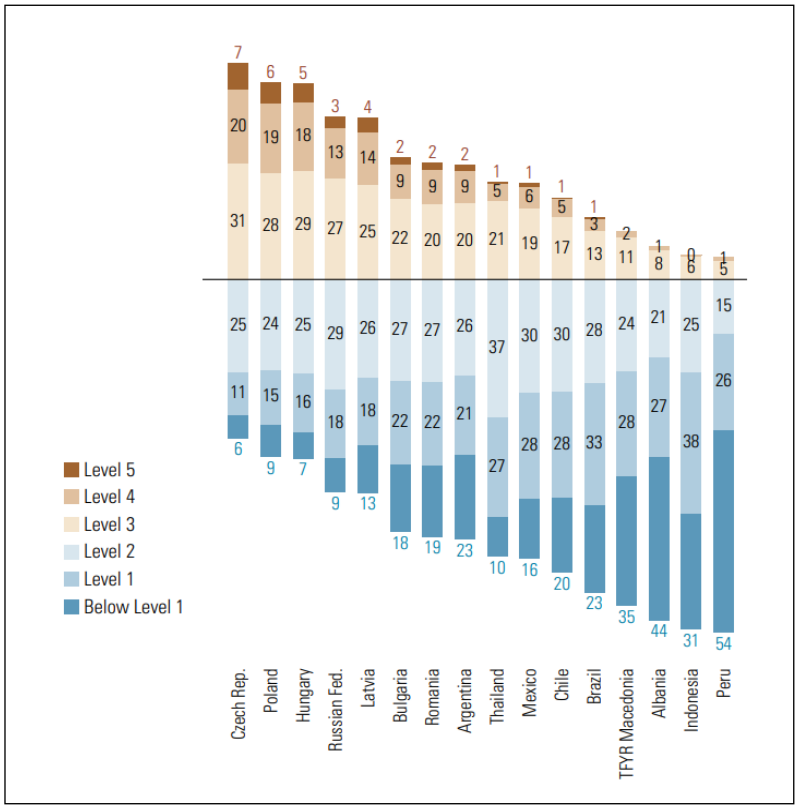
Figure 2-5. Indonesian student achievement progress

UNESCO (2004) also reported that the literacy performance of students from Indonesia's with the most favourable backgrounds is worse than OECD students with the least favourable backgrounds. This suggests that the Indonesian school system performance is problematic and unsatisfactory. When compared to middle and lower-income countries, about 69% of Indonesian students performed at or below Level 1 (Figure 2-7). Level 1 shows poor reading comprehension.

2.4.2 Quality of assessment

To assess student achievement across the nation, the MOEC employs the National Exam (Ujian Nasional [UN]: Ministry of Education and Culture, 2017b). According to the Indonesia Ministry of Education Regulation (2017), the test takes place at the end-of-year 9 and 12 as a prerequisite for a higher school level. The aim is to assess the education quality of each school following national standards (Ministry of Education and Culture, 2005, 2017b). Four subjects are examined, including Bahasa Indonesia, English, Mathematics and Science.

MOEC gathers and ranks the outcomes of the examination, then uses this ranking to map school quality and to allocate financial aid for school in need.



Source: (UNESCO, 2004, p. 123)

Figure 2-6. Reading proficiency among middle- and lower-income countries (from PISA 2000-2002)

2.4.3 Quality of assessment

To assess student achievement across the nation, the MOEC employs the National Exam (*Ujian Nasional [UN]:* Ministry of Education and Culture, 2017b). According to the Indonesia Ministry of Education Regulation (2017), the test takes place at the end-of-year 9 and 12 as a prerequisite for a higher school level. The aim is to assess the education quality of each school following national standards (Ministry of Education and Culture, 2005, 2017b). Four subjects are examined, including Bahasa Indonesia, English, Mathematics and Science.

MOEC gathers and ranks the outcomes of the examination, then uses this ranking to map school quality and to allocate financial aid for school in need.

Nonetheless, there have been serious concerns regarding the validity and reliability of the tests for examining the outcomes of education. The concern is especially given to test administration and how the exam papers are protected from leaking (Ministry of Education and Culture, 2013c). Massive cheating is evidenced that the exam results may not measure learning outcomes accurately (Berkhout et al., 2019). According to Berkhout et al. (2019), there is a pressure from the local authority for schools to have a 100% pass rate, which motivates schools to cheat the system. For this reason, the author opted not to use the National Exam score in this study.

Also, the National Exam has garnered severe criticism from some educators and scholars (Badowi, 2016; Napitupulu, 2012; Suratno, 2014; Wiwoho, 2018). They blamed the national examination system for worsening education in Indonesian because the exam only measures students' cognitive ability, which tests students' memorisation of specific topics in the lessons. The enforcement of the national examination system also signals that the central government is determined to continue to control the national education in a centralised way.

2.4.4 Constantly changed curriculum

As the national leaders in Indonesia changed, schools are required to adjust to the type of citizens that the leaders think best to serve the country. It is difficult to judge whether schools are organised to serve politicians or students (Bjork, 2004). In this section, the author shall explain the development of changes in the Indonesian educational curriculum.

Before the year 1999, the Indonesian educational system was centralised. The central government, through the Ministry of education, plays a leading role in education policy across the nation. Their role includes the decisions of the content, textbooks, hours of schooling, and

other public-school supplies. Also, according to Jawas (2008), there was minimal room for flexibility for the teachers in delivering the curriculum to the students.

Since the mid-2000s, Indonesia has employed a broad range of education reforms, decentralisation, and enhancements in teacher training standards (OECD & ADB, 2015; Tobias et al., 2014). As of 2009, the Indonesian government committed spending at least 20 per cent of the annual spending on education (OECD & ADB, 2015). This significant budget growth on education has allowed the government to eliminate school fees and enrich education through programmes financed from the national funds (Steer & Smith, 2015). For example, the school operating grants scheme *(Bantuan Operasional Sekolah - BOS*). Although BOS funding does not support schools to meet imposed minimum standards of operation (Sopantini, 2014). The highest proportion of the spending goes to teachers' salary (Tobias et al., 2014). Despite the impressive increase in public education spending, Indonesia still accounts for a smaller GDP share in this area (3.58%) than its neighbouring countries (World Bank, 2014c).

Besides increasing the national budget allocation, the education reform also included curriculum change. The curriculum transitioned from subject-based to competence-based, and from teacher-centred learning to student-centred learning (OECD & ADB, 2015; Tobias et al., 2014). So, the students are no longer just memorising materials, but also show how they can do things (Kwartolo, 2002; OECD & ADB, 2015). The curriculum reforms provide a robust foundation for teachers and schools in Indonesia to adapt teaching methods as needed (Ministry of Education and Culture, 2013a). It creates a more positive school and classroom climate. The current teaching-learning process is expressed in the Government Regulation Number 19/2005 Article 19 as follows (Indonesia, 2005; Ministry of Education and Culture, 2013a):

> *"The methods and technique of teaching in classrooms is performed throughout a sense that is interactive and collaborative, inspirational, fun, and challenging and stimulating, encourages and motivates students to take part actively, and offers enough room for initiatives, creativity, and autonomy following the students' abilities, preferences and physical and psychological development"*

This regulation appears to address the negative side of education and culture mentioned in the previous that most of the teaching-learning process is typically teacher-centred and students are often passive in their classroom.

Although progress has been made in the quality of education, the changes have posed significant challenges for teachers and the local authorities (Kwartolo, 2002). Large class size (i.e., a class with over 50 students at a time) and lack of teaching skills has been a major impediment in implementing the learner-centred approach (UNESCO, 2004). The Ministry of Education and Culture (2013a, 2013c) also identifies that despite many studies showed a positive effect of the reforms, there are some limitations in the quality of its implementation. For example, Chang et al. (2014) found that almost no significant difference in teaching practices between teachers who had undertaken the certification process, those who have trained in student-centred teaching approaches, and those that have not been certified or have not undergone the student-centred approach.

2.4.5 The teaching-learning process: teacher-centred

Changes in the national curriculum have not resulted in changes in teaching-learning practices (Chang et al., 2014). Most teachers still conducted a teacher-centred approach to learning (Ragatz et al., 2015). A qualitative study by Faridi et al. (2016) found that teacher in English lesson had difficulties in implementing the student-centred learning approach.

In terms of classroom climate, teacher-student interaction in the classroom is mostly led by teachers and typically involving lecture-style instruction (Ragatz et al., 2015). Ragatz and colleagues also found that this teacher-lead interaction made up 74% of interactions, while teacher-student interaction accounted for only 11%, and 15% are student-student interaction. The proportion of time used for group work was also low. This result is consistent with the study conducted by Maulana et al. (2011) which found that Indonesian teachers maintain their dominant behaviour. In another study, Liem et al. (2009) also found similar findings, together

with Singaporean and Filipino, Indonesia students have a lower preference to disagree with their teacher compared to Australian students. Those three South-East Asian students also had higher conformity in the classroom (Liem et al., 2009, 2016). In short, Indonesian students are passive in the classroom, and the teaching-learning process is usually dominated by the teacher.

In term of promoting critical thinking, the time spent on non-routine problem solving decreased from 12% to 4%. Ragatz et al. (2015) argued that less focus on critical thinking may be affected by the main orientation toward the National Examination. The exam has traditionally endorsed the use of memorisation and routine problem-solving. As such, the focus of education in Indonesia is at preparing students to pass examinations and get high exams' score (Berkhout et al., 2019; Effendi & Suyudi, 2016; Furaidah et al., 2015; Ragatz et al., 2015). This exam orientation might lead to an exam-oriented system of education where the teacher teaches subject to help their students pass the exam (Jennings & Bearak, 2014; Phelps, 2011).

2.4.6 Decentralisation: large quality gap across the nation

Along with the curriculum reform, the Indonesia education system is now decentralised. Major decision-making powers are passed to individual schools (OECD & ABD, 2015). It allows schools to determine their lesson plans substantially – including teaching and learning plan and preparation, learning loads, vision, timetabling and developing school curriculum – but they remain within the national regulatory guidelines (Sumintono, 2006; Tobias et al., 2014). This policy affects the mechanism on how the government delivers service for education around the country, and initially, it was aimed at giving more interest in local education management to schools and community members (Tobias et al., 2014). Intrinsically, the decentralisation approach allows a school to establish a management system that guarantees teaching and education quality (Sofo et al., 2006). However, it has not been applied effectively

through schools in Indonesia. Many Indonesian headteachers find it difficult to apply the new reform in school management because they lack leadership skills (Al-Samarrai, 2013a; Bjork, 2004, 2005; Jawas, 2008). Besides, Bjork (2004) argued that a clear lack of leadership ability was found in the district government and not just at school levels. He added that as a responsible body to run the regulation; MOEC has not yet started restructuring the culture of the education system. As pointed out by Bjork (2004), the introduction into the recent changes "has succeeded in reforming discourse, but not practice" (p. 260).

Similarly, Al-Samarrai and colleagues (2013b) conducted an evaluation survey of education quality in 50 Indonesian districts between 2009 and 2012. They found that the capability of local governments to provide quality of primary education services differs considerably across Indonesia. He argued that the ability of local governments to manage their schooling systems effectively is crucial to develop educational excellence. Decentralisation has put local governments as the principal actor in providing basic education service, predominantly district governments (Al-Samarrai, 2013b; Ministry of Education and Culture, 2013a).

Another issue is a sizeable gap of achievement in the rural, urban, and remote areas and between the Eastern and the Western part of the country (Samosir, 2008). Yeom et al., 2002) noted some problems with the decentralisation and its implementation. Such as the large disparity between provinces and districts, lack of adequate teacher training, the passive teacher approaches to policy and resource and financial insufficiency. Besides, the ability for each local government to manage and provide educational services efficiently is not equal between one district to another (Al-Samarrai, 2013b; Ministry of Education and Culture, 2013a; OECD & ADB, 2015).

In summary, the quality of education in Indonesia depends on the local government's commitment and ability in providing a superior quality of education in their districts. Moreover,

the dualism of responsibility made it more complicated. This means that the TIMSS 2011 data may not portray the huge difference across the region. TIMSS 2011 only comprises 5795 sample within 153 schools, but it still can provide tentative information and can be a basis to conduct further research.

2.4.7 Improving teacher quality

Teaching quality is a critical factor in student learning outcomes. In 2005, a new law was introduced, called Teacher and Lecturer Law. This new law acts as a starting point for the teacher certification process in Indonesia. The goal is to provide constitutional protections for the professional teacher force and to incorporate minimum education levels with opportunities for remuneration. The certification of a teacher is a consensus tool to achieve both the welfare and quality of teachers (Chang et al., 2014; Jalal et al., 2009). The government also increases the education budget, and the most substantial proportion goes to recruiting more teachers and to increase teachers' salaries (Bima & Yusrina, 2018; Tobias et al., 2014).

However, wage rises alone will not promote the academic performance of students (De Ree et al., 2017). Chang et al. (2014) noted that there are several conditions related to this situation. First, teachers have a lower educational degree than as expected in the regulation. Second, they have low topic awareness, pedagogical ability and intellectual potential. Third, teachers' pedagogical approaches are frequently ineffective and insufficient. Finally, the lack of teaching commitment and initiative among the teachers (e.g., the frequency of teacher absenteeism remains high).

Many studies have shown that the certification policy had no significant impact on teachers' teaching quality (Al-Samarrai, 2013; Chang et al., 2014; Kusumawardhani, 2017; Syahril, 2016). Indonesian students still have poor performance in several international assessments of education. The assumption that teacher certification policy would motivate the teacher to improve their teaching quality does not appear to be true (Syahril, 2016). Although

the TIMSS 2011 data used in this study seems outdated, it can at least provide a brief picture and explanation of school and classroom climate in Indonesia.

2.5 Indonesia's *Madrasah* in brief

The word *madrasah* originated from the Arabic language. *Madrasah* means a place of learning (Ma'zumi & Jakaria, 2012). A *madrasah* is an Islamic educational institution that offers training to its students to learn core Islamic subjects which typically includes, Islamic theology (*aqidah*), the *Quran*, *hadith*, jurisprudence (f*iqh*) or Islamic law, Islamic history, and Arabic language (Ministry of Religious Affairs, 2014a; Tan, 2014b). *Madrasah* is often applied to Islamic religious schools in Pakistan and Bangladesh. (Gent, 2012; Park & Niyozov, 2008). In the Indonesian context, the term *madrasah* refers to Islamic schools at the basic and secondary levels that are managed and supervised by the Ministry of Religious Affairs.

This definition must be clearly identified because, in the Indonesian context, Islamic school takes many forms (Afrianty, Hefner, & Azra, 2007; Rabasa, 2005; Tan, 2014b). Tan (2014a) acknowledged at least three forms of Islamic school, *pesantren*, *Sekolah Islam*, and *madrasah* (see Table 22). *Pesantren* is an Islamic boarding school that mainly to teach extensive and more profound Islamic studies and preparing the student to be Islamic preachers called *ulama* (Afrianty et al., 2007; Azra, 2003; Tan, 2014b). Tan (2014a) also stated that *pesantren* mainly focus on the teaching of the classical Islamic books (called yellow books – *kitab kuning*) from the leading scholars of the early Islamic era. *Pesantren* are managed and mostly operated by one or more religious person called *ulama* or *kyai* (Afrianty et al., 2007; Rabasa, 2005). Most *pesantren* are affiliated with mass Islamic organisation, such as the *Muhammadiyah* and *Nahdlatul Ulama* (Hasan & Jihad, 2008; Rabasa, 2005).

Sekolah Islam refers to any school from primary to secondary schools that embrace Islamic values and principles (Afrianty et al., 2007; Azra, 2003; Bryner, 2013; Tan, 2014b). *Sekolah Islam* is managed and supervised by the Ministry of Education and Culture, not by the

Ministry of Religious Affairs. It only delivers the same curriculum as the general school. Historically, *Sekolah Islam* was started by modern Islamist, it adopts the Western educational system without breaching Islamic values (Rosyad, 2007: Tan, 2014b).

Finally, a *madrasah* refers to primary and secondary Islamic schools, where Islamic subjects (MORA curriculum) are taught together with general subjects (MOEC curriculum), but under the supervision of MORA. *Madrasah* aims to create learners, such as those from western 'secular' schools, however, is distinguishable for fostering a greater understanding of Islam (Hasan & Jihad 2008). In this study, *madrasah* is the focus, not *Sekolah Islam*, or pesantren.

Table 2-2: Islamic school forms in Indonesia

Types	Governing body	Curriculum (created by)	Aims
Pesantren	Islamic organisation (i.e., *Nahdlatul Ulama, Muhammadiya*h) or independent religious scholar (*ulama/kyai*)	Religious scholar (*Ulama/kyai*)	To teach extensive and more profound Islamic studies and preparing the student to be Islamic preachers
Sekolah Islam (Islamic School)	MOEC	MOEC	To deliver general knowledge/curriculum and embraces Islamic values and principles to its students
Madrasah	MORA	MOEC and MORA	To teach a student who has general knowledge but having a better understanding of Islam

As mentioned in Section 2.3., most *madrasahs* are private (Ministry of Education and Culture, 2017a). The private *madrasah* often offers low-quality education (Afrianty et al., 2007; Parker & Raihani, 2009). This low-quality is because of several conditions. The private *madrasah* mostly operates in the rural area and serve the economically disadvantaged community and have limited funding (Afrianty et al., 2007). Private *madrasah* has inadequate infrastructure and facilities, lower teacher salaries, and fewer learning materials compared to the general school (ADB, 2014; Ali et al., 2011; Asadullah, 2018; ADB, 2015). Private *madrasah* also struggles to provide adequate teaching and learning environment (ADB, 2014,

2015). Student achievement in private *madrasah* is lower than in the general schools (Parker & Raihani, 2009). In contrast, public *madrasah* and the public general schools were shown to have similar levels of achievement (Newhouse & Beegle, 2006).

Moreover, with their limited resources, the students have more subjects to learn compared to students at the general school (Afrianty et al., 2007; Azra, 2015; Tan, 2014a; Zuhdi, 2006). The student learns all subjects as in the general school, as required by MOEC, and concurrently follows the curriculum designed by MORA (Azra, 2015).

With the uniqueness of *madrasah* compared to general schools, perhaps the school and classroom climate, and their effectiveness in promoting student learning outcomes are different compared to the general school. This research, therefore, also considered this different pathway of schooling in the Indonesian context.

2.6 Chapter summary

This chapter briefly explains the research context, discussing the overview of Indonesia as a country and its culture. The chapter also provides information about Indonesia Education System, the quality of education in Indonesia, and the effort that the Indonesian Government put to improve its educational quality. The chapter also explains *madrasah* as an alternative type of school to the general school.

Indonesia's educational system is reflected by its culture as a nation. Indonesians are collectivist and tend to be high in terms of power-distance. Such cultural characteristics can also be observed in the teaching-learning process. Students are expected to be trained “how to do” and have a tendency to accept one viewpoint to answer a question. Typically, students always agree with their teachers. They avoid discussions and debates in class. They follow the teacher’s orders and instructions. It should be noted that the cultural explanation in this chapter is based on Hofstede’s (1986), which is aimed to provide a brief understanding only, and not meant to describe the culture of Indonesia as a whole.

In term of the challenge to achieve its educational quality, Indonesia is struggling to improve its learning outcome quality compare to other developing countries. Political agenda also hinders the progress in improvement. The constant changes in curriculum are often a change on paper, but not in the reality. Such constant changes further hamper the progress in achieving a better quality of education.

Moreover, the quality of assessment also faces a serious concern in term of its reliability to measure student achievement progress. Massive cheating and the pressure from the local authority to have a 100% pass rate are some core issues that need to be dealt with. There are also gaps in the quality of education across the nation. Such as the gaps in terms of the quality of the teacher, financial and material resources, and leadership skills, and the local government's ability to run the school system in their respective region.

After putting the research in its context, the following chapter describes the theoretical background of the study. The researcher shall review previous research and the key concepts of this study. The literature review is presented to gain understandings of the research aims, objectives, and the research questions proposed in the previous chapter.

Chapter 3: Literature review

3.1 Introduction to the chapter

To answer the first and second research objectives, the author will review two areas of related literature. First, the literature on the history and development of school effectiveness. This review will outline the key frameworks that made school climate an important process factor for school success. Second, the author will discuss school climate as a concept as well as its development.

3.2 A brief history of school effectiveness research (SER)

Some common questions have arisen within school effectiveness research (SER). Of the key questions are: What makes one school more successful than another? Why do schools have different student learning outcomes? Several variables may be attributed to these questions, such as school intake, teaching quality, school leadership, or school facilities. However, the answers to these questions are not simple and straightforward. A diverse point of view can be used to answer the questions, such as sociology, psychology, education, and economics (Kyriakides & Creemers, 2008; Reynold et al., 2014; Scheerens, 2014). A more specific theoretical framework keen on the questions is school effectiveness/educational research (EER/SER).

The key effort of early EER/SER is to identify the fundamental features that support and explain students' learning outcomes. It is unclear to declare the precise beginning of SER (Creemers & Kyriakides, 2008; Gray, 1996; Reynolds et al., 2014; Rutter & Maughan, 2002; Sammons, 1999; Scheerens, 2014). However, Reynolds et al. (2014) summarised its history into five main phases as described in the following sections.

3.2.1 First phase

The first phase started as a response to the classic and alarming research of Coleman et al. (1966) and Jencks (1972). This is considered as alarming research because of the conclusion they made degrades the function of a school. Their studies are also considered as a fundamental point in the history of SER (Reynold et al., 2014). They showed that only a low and marginal percentage of student academic achievement can be credited to schooling compared to the effects of students' innate ability and their family backgrounds. Their conclusions led to a common belief that schools do almost nothing or schools do not make a difference and that education cannot counteract for society (Bernstein, 1970).

The apathetic conclusion was a direct challenge to the earlier presumption that choosing between schools and school process are central to the future of children (Stringfield & Teddlie, 2011). This antithesis was a critical catalyst for many scholars who began to conduct some countering research. In the next study, a better constructive understanding of the school's role was highlighted by stressing the importance of school and its impact on learning outcomes of students. For example, Weber (1971) who conducted a relatively small study reported that the poor reading ability of students with economic disadvantages was a failure of the school, not of the students or their social background. At the end of the 1970s, critical studies of school effectiveness from Brookover et al. (1978) were also confirmed Weber's conclusions. The latter studies also give more attention to the learning climate.

Another study on school improvement conducted by Edmonds (1979) also countered the pessimistic perspective of schooling. Then more extensive studies by other researchers followed (for instance: Mortimore, 1988; Reynolds, 1982; Smith & Tomlinson, 1990). All studies have found consistent linkages between the impact of schools and stressed that school has a major effect on students' achievement.

3.2.1.1 Second phase

The second phase was initiated by the using multilevel modelling approach in the mid-1980s (Goldstein, 1997; Goldstein et al., 1993; Goldstein & Woodhouse, 2000). These methodologically robust studies have begun to demonstrate the empirical consistency of school effects in the fields of study, such as the relatively stable impact of the school over time. These fields also include stability on various academic achievement, differential effects on students with varied demographic characteristics, school sizes and long-term school effects (Reynolds et al., 2014). As a result of using a multi-level approach, researchers may also compare different countries, regions and nations. The general aim of this model is to take into account regional or national policies that affect school policies and practices and that also help influence teaching and learning activities in the classroom.

3.2.1.2 Third phase

In the early 1990s, the third phase of SER began. There have been several attempts to explain why schools have diverse impacts. It is the change from "input/output" only to input-process-output and context research (Scheerens et al., 2003; Reynolds et al., 2014). Process factors provide potentials explanation as to why schools vary in outcomes (Scheerens, 1990). The process factors are meant to characterise what occurred in the school, in term of delivering teaching and learning. Some influential studies that considered process factors were research conducted by Teddlie and Stringfield (1993) in the US, the Louisiana School Effectiveness Studies. Another work was in the UK, studied subject department effects on student performance and school effects (Thomas et al., 1997). During these years, several leading reviews also appeared to the field such as by Reynolds et al. (1996), Scheerens and Bosker (1997), and Teddlie and Reynolds (2000).

This input – process – output model (Teddlie & Stringfield, 1993) is a basic model of another new model as emerging in the later phases of SER. To illustrate this model, Figure 3.1 shows examples of each part.

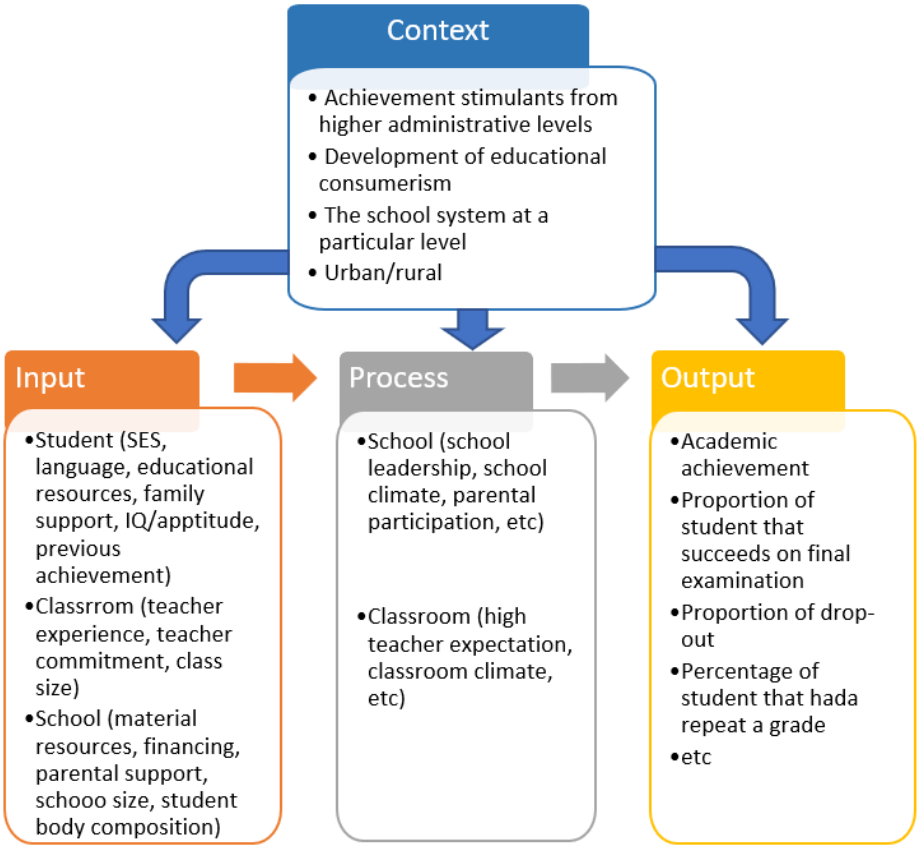


Figure 3-1. Input process output model (Scheerens et al., 2003)

This study, which aimed to look at school and classroom climate influence on student learning outcomes, therefore mainly looked at the process occurred in the school. There is some researcher (e.g., Scheerens & Bosker, 1997; Creemers & Reezigt, 1999; Scheerens et al., 2003) in this phase who stressed the importance of school and classroom climate. For example, Creemers (1994) and Creemers & Reezigt (1999) noticed the significant effect on school and classroom climate. Scheerens et al. (2003), who proposed 14 effectiveness enhancing factors which also stressed the importance of school and classroom climate. Those 14 factors are a collection of empirically supported factors, that have been shown to positively correlate to students' academic achievement in educational effectiveness study. The factors are

achievement orientation, leadership, staff's collaboration and cohesion, the quality of curriculum, school climate, potential for assessment, parental engagement, classroom climate, effective classroom management, well planned instruction, self-regulating learning, differentiation and adaptive learning instruction, and maintaining students' progress records, giving appropriate feedback and positive reinforcement.

However, in terms of inputs and process variables, some other conceptual fluidity exists (Reynolds et al., 1996). For example, the teaching qualities could be operationalised as criteria for teachers, which is viewed as input factors. However, as an alternative, when working in the context of observed teaching activities, it can be seen as a process factor.

3.2.1.3 Fourth phase

According to Reynolds et al. (2014), the fourth phase is yet in evidence currently. This phase shows the noticeable internationalisation of SER. It reflects the noticeable internationalisation of the area, combined with convergence and collaboration of methods provided by SER researchers working closely with practitioners and policymakers. A mixture of school effectiveness (SE) and school improvement (SI) contributes to a more cultural view of schooling than of the strict structural factors of SE and its determination to treat teachers as something other than pure empirical academic actors (Reynolds et al., 2014).

In this stage, the theoretical foundations and theoretical models were increasingly being developed. Concerning research methods, mixed-method studies involving large-scale quantitative exploration were initiated in conjunction with an in-depth case study of individual schools.

3.2.1.4 Fifth phase

The typical features of this phase are the establishment of SER as a complex and dynamic group of relations between effectiveness factors (Reynold et al., 2014). The most prominent movement at this stage is marked by the emergence of the dynamic model of seeing

educational effectiveness by Creemers and Kyriakides (2008). This dynamic model tries to establish a robust conception of how the system of education functions. This ambitious model (Sammons, 2009) is moving beyond associations. It attempts to discover causality in the educational system, shifts from the concept of effectiveness as a static function, positioning change assessment as one of its main goals. In turn, this more dynamic viewpoint is related to involvement in alternative types of statistical analyses (i.e., multilevel structural equation modelling) which can enable the development of specific relationships among educational factors and the outcomes of students as well as reciprocal relations between educational factors (Goldstein, 1997; Goldstein et al., 1993; Reynolds et al., 2014).

3.2.2 Critics on SER

School Effectiveness Research (SER) has a long history (as mentioned above) and have played a significant role in the development of educational research but is also contentious (Lauder et al., 1998; Thrupp, 2001). This section addresses together with a variety of counterarguments toward SER. Teddlie and Reynolds (2001) in their attempt to encounter SER's criticisers, the critics can be classified into three main issues which include: political, methodological, and theoretical.

Concerning the political issue, criticisms contend that the SER study is of "political interest," complies to government values, and supports the perception that education fails, "blaming' schools and their teachers (Elliot, 1996). Luyten et al. (2005) added that the so-called 'political-ideological aspect' of the study into school efficacy contributes to an investigative agenda that represents policy concerns instead of empirical science. However, Teddlie and Reynolds (2000) stressed that SER is always politically debatable, as it refers to the nature and goals of learning. Regardless of the position, it is evident that there are limits that are easily recognised in the field.

Thrupp (2001), Luyten et al. (2005), and Angus (1993) have addressed the critics on methodological approach in SER. They criticise that the methodological approach in SER mainly focuses on the quantitative and the excessive use of statistical analysis, compared to the use of qualitative approaches. Morley and Rassool (2002) added that learning outcomes in SER are placed within a framework of technical rationality, which represents the discourse of failure and a performance obsession as well as too “mechanistic” (Elliott, 1996). The obsession on academic performance in SER would lead a problematic side-effect, the culture of assessment-oriented teaching where the teachers teach for examination only (Dorling & Tomlinson, 2016; Jennings & Bearak, 2014; Marshall, 2017; Phelps, 2011). In the Indonesian context, the use of academic performance only in SER also appear. For example, research conducted by Kaluge (1998, 2018) as well as Creemers (1999). They mainly use academic attainment only in the research as well as only use quantitative approach. However, the SER, in fact also use different research methods, for example, by using mixed-methods and suggested the use of it (e.g. Luyten et al., 2005; Muñoz-Chereau, 2013; Sammons, 2010; Sammons et al., 1998). This research is also in the position to use not only academic achievement and strongly agree to use a combined quantitative and qualitative approach.

Regarding of the theoretical framework, Luyten et al. (2005) argued that the SER includes layers, i.e., students who are nested in classrooms or teachers, nested inside divisions/departments that are seen as subsystems in schools. Luyten et al. (2005) argue that SER neglects other things that are considered to be relevant for learning. Moreover, Luyten et al. (2005) argued that work in SER seems to have the "objectivity" to generate scientific knowledge by rigorous quantitative methodologies. Those authors claimed that SER ignores theoretical foundations in that quantitative processes rather than conceptual/theoretical explanations support the use of certain variables. This accused the SER of using the variables to pick and operationalise, but without clarifying how and why they refer to each other. There

are some SER researchers (e.g., Coates, 2003 – Microeconomic theory; Kyriakides et al., 2000 – Cremers comprehensive model; Reezigt et al., 1999 – Carrol model) that have used the theoretical framework in their research, even if not all of them (Scheerens, 2013).

3.2.3 Choosing among the various models of SER

This study utilised the dynamic model of educational effectiveness as its conceptual framework. This framework is selected over the others due to two main reasons. First, the model considers a broader range of educational outcomes, including effective outcomes such as academic self-concept and self-efficacy that will be used in this study. Second, the model considers five effectiveness evaluation dimensions (frequency, focus, stage, quality and differentiation) in explaining factors that influence school effectiveness. It enables researchers to explain in depth why some schools have a better climate than others. The utilisation of this model in Indonesia is essential considering the diversity of educational quality in the country. This model enables researchers to highlight differences among the schools in greater clarity.

There are, of course, other approaches to study school effectiveness. From an organisational point of view, school effectiveness is seen as a function of effective policies and management. From this point of view, school effectiveness may be promoted, for instance, through a change in curriculum. But such a perspective often neglects the psychological components of school effectiveness. After all, schools consist of individuals with their respective roles, obligations, motivation and goals. Such collections of individuals have unique experiences that may contribute to the development of the effectiveness of their school. As will be discussed in detail, DMEE allows researchers to assess both the organisational components of schools as well as the psychological experiences of school members.

3.2.4 Dynamic models of educational effectiveness (DMEE)

3.2.4.1 Background

Since the primary purposes for choosing this framework lies on two purposes as mentioned in the above section, a detailed explanation of DMEE goes beyond the extent of this section. Therefore, this section presents a brief overview of the dynamic theory in brief as well as the concept of effectiveness evaluation dimensions used in the current study. For a more extensive overview, please refer to the book *The Dynamics of Educational Effectiveness* (Creemers & Kyriakides, 2008).

In response to many criticisms on school effectiveness research as mentioned in previous section (see Angus, 1993; Elliott, 1996; Goldstein & Woodhouse, 2000; Luyten et al., 2005; Thrupp, 2001), Creemers and Kyriakides (2008) develop the DMEE. The model aimed to help and expand practice throughout a theory-driven and an evidence-based approach. Also, DMEE develops new models that move beyond the estimation of statistical relationships between variables (Coe & Fitz Gibbon, 1998). It stands out as an alternate theoretic model that counters the methodological and conceptual perspectives present in the field. DMEE has been not just remarked as an enhancement in the SER theoretical field and methodology by many researchers (e.g., Reynolds et al., 2014; Sammons, 2009). However, it has also progressively surpassed earlier models. It has been so dominant, that it has been highlighted as the starting point of the phase 5 of SER, which reflects of educational effectiveness as a dynamic, multilevel, not static set of relationships and achieving different outcomes. It is, therefore, a far more complex conceptualisation of the relationship between educational factors than having been used in SER formerly.

The model builds on the Creemers' (1994) comprehensive model of educational effectiveness. The model is relatively similar to other models, such as multilevel school's effectiveness models developed by Scheerens (1992), Slater and Teddlie (1992), and Teddlie

and Stringfield (1993). Models also shared characteristics in integrating school and classroom level variables that influence student learning outcomes.

The key idea of the dynamic model assumes that effective schooling as an ongoing dynamic process. It attempts to explain the complex relations between the numerous factors of school effectiveness and education development factors (Sammons, 2009). The model tested by several studies in some different contexts, but mostly in Cyprus and Netherlands (Kyriakides & Creemers, 2009, 2012; Kyriakides et al., 2010; Kyriakides et al., 2013). The model refers to various factors of effectiveness which operate at different levels. It has four levels: context/national policy, school, teacher/classroom, and student. Teaching and learning process at teacher/classroom level is also emphasised. However, school-level factors are expected to provide the conditions for maximising the teaching-learning process by creating school policy on teaching and providing a supportive learning environment.

The model also considers that the teaching-learning process is influenced by a broader educational setting. For example, social values and national policies play a significant role in determining teachers' and students' expectations. According to Creemers and Kyriakides (2006, 2008), the model presumes that the factors at school level and context level have effects on student learning outcomes both directly and indirectly. Also, this model considers a broader range of educational outcomes. Also, some factors which operate at the same level are assumed to be interrelated and not always linear. The presumption offers the opportunity to search at acceptable values of the various dimensions of factors and for maximum combinations of factors. A major feature of this model also changes in features and effectiveness over time; hence the use of the term dynamic (Creemers & Kyriakides, 2006, 2008).

3.2.4.2 Effectiveness evaluation dimensions

To counter another critique that the models of educational effectiveness generally fail to adequately provide effectiveness measure factors satisfactorily. The model is designed to

measure and describe school effectiveness factors using five evaluation dimensions, namely: frequency, focus, stage, quality and differentiation (Creemers & Kyriakides, 2006, 2008). Those effectiveness evaluative dimensions allow researchers to explain more precisely the functions of each school effectiveness factors. Generally, the *frequency* dimension represents the number of activities related to the factors of effectiveness, while the other dimensions evaluate the quality of characteristics of each school effectiveness factors. This may be the most straightforward way to assess the school effectiveness factors that influence student performance, and most effective research use that same dimension to identify factors of effectiveness (Creemers & Kyriakides, 2008). By only using frequency, many practices may not necessarily improve student learning outcomes. It is therefore important to evaluate their quality regularly, for example, by taking a critical look at the specified activities and examining whether the empirical literature fully supports them (Creemers & Kyriakides, 2006, 2008).

The focus dimension refers to the broadness or specificity characteristics of activities: whether it is too broad or too specific related to its purpose. *The stage* dimension refers to the specific moment when activity happens. *The quality* dimension associated with activities or policies that have been applied to make sure that the factors are functioning as they are. Lastly, *the differentiation* dimension applies to the degree to which factor-related activities for all subjects concerned are performed in the same manner.

3.2.4.3 Factors operating at different levels

3.2.4.3.1 Student Level

Previous research on SER consistently found that student achievement is influenced by students' characteristics/background. Therefore, in the dynamic model, this student-level characteristic is also emphasised. Further, the dynamic model includes two key categories of student's background factors that are taken from a different perspective of SER; sociology and psychology. From a sociological standpoint, the model includes student's socio-cultural and

economic backgrounds, including Social Economic Status (SES), ethnic background, and gender. On the other hand, from the psychological perspective of SER, student's characteristics are also considered (Creemers & Kyriakides, 2008). These characteristics include aptitude, motivation (perseverance, subject motivation), and personal characteristics (personality traits and thinking style). Other variables are time on task and opportunity to learn. However, in term of motivation variables, it is problematic. Creemers and Kyriakides (2008) did not include self-efficacy, even though this variable in many studies have been found to have a relationship with achievement (Marsh & Martin, 2011; Pajares & Urdan, 2002; Usher & Pajares, 2009). Therefore, in this research, the researcher includes motivation variables like self-efficacy as well as self-concept, particularly math self-efficacy and self-concept.

3.2.4.3.2 Teacher/classroom level

Since the main purpose of SER is to establish specific factors that support and promote learning, Creemers and Kyriakides (2008) argued that the dynamic model emphasises mostly on the factors affecting at the teacher and school levels. At the teacher or classroom level, their focus is mainly on the teaching-learning process delivered by the teacher. They proposed eight effectiveness factors that portray the pedagogical role of teacher. Therefore, even they do recognise that teacher's characteristics (i.e., gender, age, education background) are important to student learning outcomes and the main topic in SER, those factors were not included in the model. The model, as mention before, believed that the process is more important than teachers' personal background. The eight factors of "the classroom process" are orientation, structuring, questioning, teaching-modelling, applications, teacher role in making the classroom a learning environment, management of time, and classroom assessment (p. 103). The explanation of each factor will be discussed in brief to get an idea of how the model emphasised to these eight factors (Creemers & Kyriakides, 2008).

(1) *Orientation* relates to the teacher's activities which provide the specific targets of the lesson being taught. Classroom orientation aims to create a supporting learning environment as well as to build a good interaction between student and teacher in the classroom (Deemer, 2004). However, not only discussing the specific goal that will be achieved, but orientation also includes physical orientation of the classroom (i.e., a seating arrangement) (German et al., 2020; Gremmen et al., 2016; Hastings & Schwieso, 1995; Zerin, 2009). In addition, the teacher, during the lesson process, encourages the students to understand why a particular thing or activity happens, by giving guidance assignment. Therefore, teacher need to use active learning strategies to make students on the track, for example by doing structured debate (Koklanaris et al., 2008; Oros, 2007).

(2) *Structuring* refers to teacher activities that not only providing a good learning material but also organised and structured it by (a) begin by explaining the objective; (b) sketching out the content to be covered and communicate the shifting between lesson; (c) reflecting on key ideas; (4) and reviewing key ideas when ending the lesson. Structuring the classroom activities also means designing a well-managed learning environment (Renne, 1996; Tanner, 2013). Accordingly, the student and teacher know what they have to do in order to create a respectful interaction in the classroom. However, highly structured lesson plan may lead teacher to only focusing the activities rather than the learning goal (Boikhutso, 2010).

(3) *Questioning* means how effective teachers raise multiple questions and try to engage students. Asking lots of question to the student could encourage student participate in the class discussion (Kyriakides & Creemers, 2006). Questioning is a potential tool to engage student involvement and enhance critical thinking (Elliot, 1994; Eshach et al., 2014). However, the teacher needs to be clear when asking question, because asking for a vague question (i.e., asking trick and too abstract/complex questions for children of their age)

may lead to some negative learning attitudes (feeling learned helplessness, and they say: 'I can't do this,' or 'It's difficult'). In addition, Excessive questioning will also disturb the learning process and decrease the quality of the student responses (Larson & Lovelace, 2013). The vague and excessive questions hamper the development of a positive classroom environment.

(4) *Teaching-modelling* refers to teacher activities which can help students to create their own approaches and strategies to help them overcome various problems. in other word, the teacher needs to develop their student's self-regulation learning (SRL) skills, because the process of learning requires self-regulation. These strategies allow students to develop and improve learning habits (Zimmerman, 2000). However, the use of SRL strategies requests for thoughtful consideration of student abilities and needs. For example, a research by Moos and Ringdal (2012) found that students with cognitive strategies, the explicit instruction on these SRL processes may have adverse reactions.

(5) *Applications* underline explicitly the immediate practice of lessons/topic that has been learned at the individual level or in a group during the course. Practice is one of encouraging strategies to improve classroom learning (Karpicke et al., 2016; Moreira et al., 2019). Besides, immediate practice significantly has more long-term retention of the materials studied than just traditional test (Pashler et al., 2007). However, this method of teaching must be followed by giving feedback (Goossens et al., 2014).

(6) *Creating a positive classroom climate.* The dynamic model of creating positive classroom climate refers to the role of the teacher in creating a positive learning atmosphere in the classroom. As this factor is a main interest in this research, a relatively detail discussion will be explained further after this short explanation of teacher/classroom level factors. However, all other factors are related to the creating a positive classroom climate factor.

(7) *Management of time*, where teachers should arrange and manage the classroom environment as an effective learning environment and thus improve participation levels and actively assess the classrooms. Occasionally, the transition from task to task can cause class chaos, particularly when students are confused with the next task, the management of the classroom can be completely disrupted (Codding & Smyth, 2008). Therefore, time management also relate to how teacher structuring the productive classroom. Meta-analysis conducted by some researcher (Hattie, 2012; Marzano et al., 2003; Scheerens, et al., 2013) emphasised that classroom time management is a productive use, which means that the most important thing is creating conditions for engaging students, not only prolonging them.

(8) *Assessment* is known as a specific aspect of teaching. The student assessment will help teachers to decide the needs of their students and to critically assess their teaching practices. Assessment is an essential element of learning in determining whether or not the learning goals are accomplished (Earl, 2012), especially formative assessment which is one of the most important factors in terms of effectiveness at any level and particularly at the level of classrooms (Biggs, 1998; Brookhart, 2001). However, the evaluation of education can only really benefit education if it is designed to provide quality education (Felner et al., 1995).

As the focus of this research is related to school and classroom climate, at this level, teacher responsibility in creating the classroom as a learning climate is discussed. As mentioned before, that this factor cannot be separated to other factors of classroom/teacher level (see the definition of each factor above). Creating classroom climate is like a central factor in classroom management. This factor can be seen as a classroom climate created by teachers and interaction that they have been made in the classroom. In a dynamic model (Creemers & Kyriakides, 2009), there are five elements of the classroom climate that

should be considered: 1) interaction between teacher and student, 2) between student interaction, 3) how teacher treats students, 4) competition between students and 5) classroom order. As empirical research in the class environment has shown (e.g., Withal, 1949; Moos, 1980: Rubie-Davis, 2014), the first two components are key elements for classroom climate assessment. The other three elements refer to the teacher's effort to create a supportive environment for learning (Creemers & Kyriakides, 2008; Creemers & Reezigt, 1999). Creemers and Kyriakides (2008) emphasised that the dynamic model's view on classroom climate refers to the type of interactions that occur in a classroom, rather than perceiving teacher behaviour. So, the stressing point is on the process of the teaching-learning itself.

3.2.4.3.3 School level

From the dynamic model's view, the school level factors are on two main elements of school policy that influence the teaching and learning process at the teacher and student levels. The two key elements are (1) teaching learning policy and (2) school's learning environment policy. The factors related to the school policy primarily contribute to the actions taken by the school to help teachers and others understand clearly what to do. Thus, there are four key factors need to consider at the school level. (1) School policy on teaching and concrete measures to improve learning. (2) Evaluation of teaching school policies and steps taken to improve teaching. (3) Regulation to create an environment for school learning and steps taken to improve the environment for school learning. (4) assessment of the school learning environment.

The primary assumption of focusing on those factors is because the model is not focused on an individual, but the effect of the action. So instead of measuring principal leadership style or so on, the model gives more emphasis on the impact or result of the leadership (e.g., policy to improve teaching and learning process at school).

3.2.5 Limitations in using DMEE in the study

It is beyond the scope of the study to test and seek to validate DMEE in the Indonesian context. The main aim to utilise DMEE as the foundation of the framework of this study is mainly on the use of its effectiveness evaluation dimensions. It is not possible to apply all the features in the model to this research for some reasons.

1. This research only uses TIMSS data that naturally cross-sectional, although the dynamic model prerequisite longitudinal data. Therefore, the DMEE cannot fully adopted.
2. Educational effectiveness model of Creemers and Kyriakides (2008) does not provide a clear definition between the learning environment and school climate, they used the term interchangeably. The model also has a limited explanation of school and classroom climate. However, they provide a chance to assess school climate by taking into account the five evaluation dimensions of educational effectiveness (frequency, focus, stage, quality, and differentiation). By using these dimensions, it is possible to measure school climate in depth. So, it may get a better understanding of school and classroom climate in the Indonesian context.
3. Thapa et al. (2013) consistently used school climate term and did more in doing a review on school climate research. Therefore, they provided a clear understanding of the school climate. However, most literature in school climate did not offer how to measure school climate in depth. School climate is typically measured by gaining school members' perception by using questionnaires, and the result mostly captured school climate perception on the four aspects of the climate. Only providing those aspects might be not fair when comparing one school climate to others, because each school has different effectiveness, may have a different focus, stage, quality, or differentiation (Azkiyah et al., 2014; Creemers & Kyriakides, 2009, 2010; Kyriakides

& Creemers, 2009; Kyriakides et al., 2009; Kyriakides et al., 2010). Besides, the psychological tradition of classroom environment research paid much attention to instruments for the measurement of students' perceptions of climate. Many studies report on psychometric characteristics (Fisher & Fraser, 1990; Fraser & Fisher, 1986). Also, there is a need to integrate elements of different research traditions and search for the contribution of the teacher in creating the classroom as a learning environment (Creemers & Kyriakides, 2008).

By combining these concepts, it will measure the effectiveness of school climate, not just the school climate. So, the whole picture of school climate might be captured, and this approach may give more information to the policymakers how to develop and evaluate school climate and may lead to improve it.

3.3 School climate

3.3.1 Significance of school climate in SER and its critiques

3.3.1.1 Significance

Though school climate has been recognised for more than 100 years (Cohen et al., 2009), the empirical study concerning school climate was started at the end of the 1970s when SER paid particular attention to the school processes (Van Houtte, 2005). The importance of school climate in SER also has been indicated in some early research as can be seen in the first phase of SER development (Bloom, 1976; Brookover et al., 1978; Rutter, 1979). Bloom's (1976) research report indicated that the learning environment has a significant effect on the results of student learning.

Bloom (1976) also said, despite evidence of variations in school learning's existence and continuity, he remained persuaded that a large proportion of the disparity was attributed to the school learning environment. Bloom (1976) examined substantial evidence for the development of various characteristics from longitudinal studies of

school learning environments and concluded "*the environment is a determiner of the extent and kind of change taking place in a particular characteristic*" (p. 209). The five-year research conducted by Rutter et al. (1979) has shown that a variety of climate and environmental factors have an impact on academic performance. Also, Reynolds (1982) concluded that school climate appeared to vary concerning the effectiveness of the school. Dumaresq and Blust (1981) identified that the school climate is affected by school members' beliefs, cultural values, attitudes that impact the school environment's circumstances, activities and practices. They argued that the school climate is acknowledged as one of the characteristics that specify or at least influence how well schools perform. Knight (1985) also indicated the urgency of further study into the links between academic performance and climate factors if schools need to be encouraged to know about their vision and missions. Later, Freiberg and Stein (1999) claimed that the climate of a school could promote resilience or become a protective factor in a school.

Also, in the DMEE, school climate is one of the main factors at the school level, noted as a school learning environment (Creemers & Kyriakides, 2008). By reviewing various school effectiveness studies, Scheerens et al. (2003) also identified that school and classroom climate are two of fourteen factors that may enhance the effectiveness of a school.

Following the multilevel nature of the school, Creemers and Kyriakides (2008) claimed that learning environment (or school climate in this research) of the school might improve the functioning of classroom-level factors including classroom climate. The classroom climate aspect emphasises on orderliness and the relationship between teacher and student, student-student relationship, and the satisfaction of this relationship (Scheerens et al., 2003).

The school climate emerged as a substantial factor of an effective school, and a close relationship between school climate and student achievement. In other words, the school participates in promoting learning outcome through its "climate". Student outcomes, whether it is academic or affective, exist in a manner where different social environments tend to

produce a range of effects from school processes (Anderson, 1982). Also, Fisher and Fraser (1990) found out that the school environment provides a significant contribution to the effectiveness of a school, together with curricula, infrastructure and leadership.

School climate is one of school process features in SER (others like teaching practices, school management and policy) that influence student learning outcomes (Reynolds, 1982; Reynolds & Teddlie, 2002; Reynolds et al., 2015; Scheerens et al., 2003). School process is believed to have the most potential for understanding and improving school success. Though many schools, or most of schools at least, have little control of the backgrounds and traits, resources and social characteristics of students. On the other hand, school process factors have enough control over how schools are managed and organised.

Some researchers (Raudenbush & Willms, 1995; Scheerens, 1990; Willms & Raudenbush, 1989) have also mentioned the process factors as “Type B effects”. Because, when statistical adjustments are made for the effects of other factors, school process factors share a more appropriate and better basis for comparing school performances. Recently, Reynolds et al. (2015) emphasised the need to acquire a better understanding of the school process factor because research on these factors in SER is limited.

3.3.1.2 Critics on school climate research

Hoy and Hannum (1997) criticise the way in which SER investigates school climate. They argued that school climate in SER is a global construct that SER researchers “often use loosely to group together studies of the school environment, learning environment, learning climate, sense of community, leadership, academic climate, and social climate” (p. 295). The researcher supports this view since most of SER did not provide a clear definition of school climate as one of the school processes factors. The operational definition of school climate tends to have a lack of theoretical basis for the selection, more common sense and varies greatly amongst studies (Coe & Fitz-Gibbon, 1998; Miller & Fredericks, 1990; Sandoval-Hernandez,

2008). If that loose concept of school climate is found to be statistically significant, then they are regarded as critical to school effectiveness (Coe & Fitz-Gibbon, 1998; Lauder et al., 1998). According to Ryle (2015), this is a pure false construct that can be classified as a categorical mistake. This is only *reification*, which interprets an abstract concept as a concrete thing (Ryle, 2015). Reifications and ambiguous definitions are not, of course, an exclusive SER problem; they are actually a common issue in the social sciences (Billig, 2013). This common problem has been raised at least since Aristotle.

This research is different from most of SER in term of the clarification of school climate concept. Therefore, there is a need to make clear the theoretical background and concept. This research provides a theory-driven concept of school climate in the study. Some conceptual differences are therefore introduced to provide a place in which factors of the school's climate can be operationalised in a theoretical model. The following section addresses current school climate definitions and concepts.

3.3.2 School climate: a challenging definition

School climate is a topic of many researchers because of its relative importance in the success of students learning (Barclay & Wu, 1980; Faour, 2012; Fraser et al., 1988; Freiberg & Stein, 1999; Hoy et al., 2002; Knight, 1985; Marsh et al., 2012; Tagiuri, 1968; Van Horn, 2003; Van Houtte & Van Maele, 2011; Wang & Eccles, 2012). However, it is not easy to define because of the volume of study, the many variables process, concepts, and models. It is a loose concept, and there is still no consensus among researchers in identifying and selecting the aspects that contribute to constructing the school climate (Carrasco Ogaz, 2016; Cornell et al., 2016; Finlayson, 1987; Freiberg & Stein, 1999; Rudasill et al., 2018). It is maybe because ‘climate is mostly an affective or feeling element of learning’ (Freiberg, 1999, p. 10). Among the definitions contain different terms and concepts including attitudes, share perceptions, emotions, and beliefs (Carrasco Ogaz, 2016).

There is a varied range of features that has emerged under the school climate research. In the early review of school climate research, the definition of school climate has been viewed more intuitively rather than empirically (Anderson, 1982). Finlayson (1987) argued that the conceptual definition of school climate, as taken by early school climate researchers, tend to use climate as a metaphor, where they maintain the idea that climate is something 'out there' in the school environment. Winter (1987) also found that some studies (i.e. Brookover et al., 1979; Edmonds, 1979), school climate was described as a school building's atmosphere which can be sensed by the impressions, moods and feelings when someone is in some part of the school (corridors, classrooms, playground). Another example, Halpin and Croft (1963) defined school climate as the personality of a school. They stated that the climate is real and can be experienced by all school members' interactions. Tagiuri (1968) described the school environment as characteristics of the entire environment. Further, in Tagiuri's view, the school climate is a formation of relatively stable elements of the *ecology, milieu, social system,* and *culture*.

Anderson (1982) also adopted those four variables from Tagiuri (1968). Again, this formation is like personal characteristics that represent a personality. *Ecology* includes building and facilities, technology, interaction, and academic or simplified by Anderson (1982) as school physical characteristics. *Milieu* indicates to the school social system, includes teacher's skills, motivation, job satisfaction, feelings, values, student's socio-economic level, teacher's educational as well as the principal's leadership. Similarly, In Anderson's view milieu is related to the features of teacher morals and student-body. *The social system* refers to the school organisational structure that comprises communication forms, instruction, supervision, management, decision-making practices. Equally, Anderson (1982) assumed that the social system dealt with financial, economic, and staff and teacher ties. *The culture* consists of behaviour patterns, values, norms, beliefs, history, and ways of thinking. The culture also related to teacher participation, teamwork, and school targets (Anderson, 1982).

Then, Hoy (1990), defined school climate is an overall concept that portrays the atmosphere of a school. School's members felt and experienced it, including student, teachers, headteachers and administrator, which affects school behaviours and activities (Hoy & Miskel, 2013). They also use the organisational/school health metaphor to investigate the school climate. Hoy's (1990) idea was adopted from Miles (1965) who was the first scholar to introduce the concept of healthy organisation in the school context. A healthy organisation is the one that not only has surviving abilities but also has the coping ability for an extended period. It continually develops and expands its coping skills. According to Hoy et al. (2002), a healthy school climate has positive members interrelationships. For example, teachers like to work with their colleagues, school, and their students and they are inspired to achieve academic excellence.

This view (the use a metaphor to describe school climate) is challenging to be operationalised or measured, and Finlayson (1987) has warned and critiqued the use of metaphors in school climate definition. He stated that the use of metaphor is ambiguous because individuals establish their own meaning of the metaphor but all of them rely on its importance for themselves. For instance, this comment illustrates well the challenges a teacher had to convey the perceived school climate:

> *"I can define it, and yet I cannot define it. It is like a ghost: I can touch it, but it is not there, but I know it is there, and I do not have a great deal of control over it." (Finlayson, 1987. p. 163).*

Then he added that psychological perspective conceptualises the process of meaning formation at the level of the individual school members. The perspective also did not take into account the interaction, organisation or culture processes where systems of meaning used by school's members are built and these processes are sustained (Finlayson, 1987). The school members also perceived the climate of a school differ, unique and vary remarkably; these are

unique from the simplest thing, such as paint colour on the walls, the way seats are positioned in the teachers' room, the presence or absence of pupil displays, and the way school leaders interact with each other. In other words, the variability of perceptions is from physical to intangible factors.

Recently, school climate has been defined as a quality of school life, for example, Freiberg and Stein (1999) described school climate as the quality of a school that assists school members to feel emotional interest, honour and significance when helping to create a sense of belonging at the same time. Hoy and colleagues (Hoy & Hannum, 1997; Hoy & Miskel, 1996, 2005; Hoy et al., 2006) stated that school climate is a collection of internal characteristics that influences the quality of the school life.

Cohen et al. (2009) prefer to the idea of quality and character of school life, comprising the perception patterns of school members' experiences of school life, reflecting the norms, goals, values, interpersonal relations, teaching-learning process, as well as school structures. Though this definition is relatively too broad (Carrasco Ogaz, 2016), however, the definition is relatively common among the researchers. The consensus of the definition is indeed limited, but most concepts highlight the importance of relationships between the members of the school family and indeed the importance of common school goals, norms and values (Payne, 2018) as Cohen et al. (2009) proposed.

3.3.3 School culture and school climate

It is important to make it clear that the term of school climate compares to other theoretically similar terms before further school climate definitions are discussed. For example, the terms of school climate and school culture are often used interchangeably (Van Houtte, 2005; Van Houtte & Van Maele, 2011). Though they are originated from different research traditions. Both concepts detect specific features of the organisation. Numerous features of the two concepts overlap when researchers want to study school contexts in order to take into

account the essence of the feel or meaning of the school. Some research reports confused these two concepts and used them together to measure organisational environments. However, this understanding has not been accepted by scholars of culture and climate. Therefore, it is necessary to make a contrast comparison of these two concepts and draw precise positions to the theoretical base for this study though both terms were vague, and neither explained deeply (Hoy, 1990).

The difference is generally that culture includes assumptions and ideologies, the climate is described as emotional and behavioural perceptions (Dorina, 2013). Hoy (1990) distinguished that school climate is generally investigated from a framework of psychology, while school culture mostly is examined from the framework of anthropology. Therefore, school climate research mostly utilises statistical analysis (quantitative) and culture studies predominantly using phenomenologist and ethnographer's methods (qualitative) (Hoy, 1990; Schoen & Teddlie, 2008). The culture of the school can be studied through stories debates, events, teacher and student reports, interviews and videos that explain what occurs in school and classroom (Schoen & Teddlie, 2008), while school climate is mostly measured by a validated questionnaire (Freiberg & Stein, 1999). Climate is typically seen as collective perceptions, whereas culture is seen as shared assumptions (i.e., norm, traditions, beliefs) (Van Houtte, 2005; Van Houtte & Van Maele, 2011). School culture is “an unseen and an unobservable force behind school activities, a unifying theme that provides meaning, direction, and mobilisation for school member” (Prosser, 1999, p. 14). It informs members of the school, which is essential and how they are to behave.

However, the debates are still going on, for example, Van Houtte (2005) tended to subsume culture under the climate, but he suggested that school culture is the better frame for studying school effectiveness. Whereas Schoen and Teddlie (2008) suggested that school culture and school climate definitions are similar but should be defined at a different level.

Then, they recommended that the school climate is more suitable to consider as a part of the wider school culture construct. This view is in line with Tagiuri (1968), which included culture in assessing school climate (Anderson, 1982; Van Houtte, 2005). According to this view, the climate of a school may reflect the culture of a school and the shared norms and beliefs influence school members' behaviours and emotive reactions to the school and thus influence the school organisational climate.

It is then crucial to further investigate how researchers define them when a comparison of the two is made. School culture researchers are more focused on the progression of school social systems over time. The importance of a deep understanding of underlying assumptions as well as the insiders' point of view of the organisation is mostly needed. School climate researchers, however, are more concerned with the school members' perceptions of observable practices and the process of the school. These practices and processes are close to the surface of the school life. They are related to the impact that the school or organisational systems have on groups and individuals (Liu, 2004; Louis, 2006; Schoen & Teddlie, 2008).

In terms of research focus, Hoy and Tarter (1997) argued that if the research goal is to establish the fundamental forces that drive action in a school or the values and symbolism of a school, then a cultural approach is preferred. Then, if the emphasis is to clarify, control and improve the actual actions of school members, the climate approach is more appropriate.

Since this research focuses mainly about how school members perceive about their school, the term school climate will be used, then the school climate was measured by the perception of all school members (Glover & Coleman, 2005; Hoy et al., 1998; Zullig et al., 2011). The stress on perception is central to school climate in order to discriminate against school culture (Brault et al., 2014).

3.3.4 Classroom climate and school climate

A classroom is a central place for student personal and academic development, and each class has a distinct climate that facilitates growth (Moos, 1980). The classroom climate is suggested as a dynamic social system which not only includes teacher-student behaviour and interaction but also student-student interaction (Moos, 1980). Previously, Withall (1949, 1969) argued that classroom climate is the emotive tone of direct interactions happens in the classroom. Similarly, Anderson et al. (2004) stated that classroom climate is the social environment of a classroom. Moos (1980) defined the classroom climate as the people's shared perceptions in that classroom.

The stressing on social interaction is one of the central 'process indicators' associated with school climate. In line with this view, in the SER, classroom interaction is also fundamental in studying classroom climate. For example, DMEE defines five foundations of the classroom as a learning environment that are needed to consider, including teacher-student interaction and student-student interaction (Creemers & Kyriakides, 2008). Those interactions are a key point since it creates how classroom members perceived the climate. Whether it is positive or negative, supporting or discourage. Scheerens et al. (2003) in their summary of 14 effectiveness enhancing factors consider classroom climate as one of the key elements and also put teacher-student and student-student interaction as central measures. Rubie-Davies (2007, 2014) also has a related view. He stated that the climate in classrooms is the product of two facets: the teaching practices and psychological matters. The first demonstrates the outcomes of teaching choices made by teachers while the second is the reflection of the interaction between students and teachers and students. Thus, classroom climate can be stated as part of the school climate that occurred inside the classroom, the teaching-learning process. The teaching-learning process itself automatically includes interactions.

Further, Moos (1980) constructed a classroom climate scale that mainly consists of relationship/interaction, self-development, and system maintenance and change factors. The relationship factors measure student involvement, support and help between student, companionship/friendship and trustworthiness, and teacher's support to the students. Personal growth measures the emphasis on the academic task. Lastly, system maintenance and change associated with maintaining the classroom functioning in order. Again, Moos (1980) also support the view that interaction is a key factor in classroom climate.

Previous studies found that classrooms climate had a significant relationship with student outcomes both cognitive and affective in some countries like in the US, China, and the Netherlands, as well as in Indonesia (Byrne, 1984; Cheng, 1994; Creemers & Reezigt, 1999; Fisher & Fraser, 1983; Margianti et al., 2001). However, most of the study was mostly using only quantitative approach (Barclay & Wu, 1980; Lin & Crawley, 1987; Moos, 1980; Reyes et al., 2012; Teodorović, 2011).

3.3.5 Issues in school climate research

Some researcher found that, there is a series problem that is essential to be focused and fixed in school climate research (Anderson, 1982; Cohen et al., 2009; Zullig et al., 2010; Freiberg, 1998; Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015). However, that is not always straightforward and there is no precise way to address such issues. To overcome the issues, Carrasco Ogaz (2016) classified eight specific challenges underlying school climate research, there are: multidimensionality, nesting or hierarchical issues, inference level, reference level, complex relationships, methodological bias, causality and temporality. However, this research only discusses two main issues because it is related to the recent study in term of to overcome conceptual clarity of school climate factors. The two issues are 1) multidimensionality, and 2) inference level and referent issues.

3.3.5.1 Multidimensionality

Carrasco Ogaz (2016) argued that the major problem that lies following the school climate multidimensionality issue is the interpretation of its concept into a complete construct view. One common problem in this field is by assuming a concept as a construct, and this can lead to *categorical mistake* (Holth, 2001; Ryle, 2015). *Categorical mistake* is the failure to assign a quality or act to something which can only be attributed correctly to objects in other categories, for instance, treating abstract concepts as if they have a physical state (Stevenson, 2010). Therefore, it is essential to separate constructs from concepts (Markus, 2008). Constructs are theory-driven and can be taken into account in the evaluation of the observed disparity between subjects (Cronbach & Meehl, 1955); whereas the concept is "an abstraction that describes a portion of reality" (Shoemaker et al., 2004, p. 15) do not have such a weight. Therefore, school climate should be seen as an umbrella concept (Freiberg & Stein, 1999). School climate factors from different research can be classified into different constructs. This approach facilitates to explain how school climate construct seems unclear and ambiguous.

For instance, when a research result state that 'positive school climate is a key component of effective schools', it is not an interpretable claim for that school climate since the readers do not know exactly which of the possible factors the researcher used. Nonetheless, the condition will be different if the school climate is seen as an umbrella concept which deal with various constructs. Therefore, "rather than trying to establish a static definition" (Freiberg & Stein, 1999, p. 28)**,** by functioning this idea, it can interpret the claim into a different context, for instance 'positive features of the school environment are a key element of effective schools. It can be construed as a more concrete idea as 'higher levels of the teacher-student relationship are positively related to schools' academic success' (Carrasco Ogaz, 2016).

3.3.5.2 Inference level and referent.

The nested design of the school climate measure is another common problem within school climate research (Marsh et al., 2012; Wang & Degol, 2016), particularly for measures that obtained from specific respondents in classrooms or schools. The problem is caused by the difference between measurements from the same cluster. Choosing to ignore this data structure could cause researchers to commit *Type I mistakes* which is endorsing results that are not real. Accordingly, it leads to misplace the level of conclusion and to make inference at the inappropriate level. Interpreting student-level perceptions of climate as though they reflect school or classroom level climate is a typical sample of the *ecological mistake* (Morin et al., 2013). Contrary, the *atomistic mistake* arises when connections between findings at an aggregate level or cluster are believed to be replicated. In short, the nested structure of the information from school climate measurements must be accurately calculated for data dependency within clusters.

Regarding the inference level and referent, some researchers claimed that school climate is a school-level variable (James, 1982; Marsh et al., 2012; Morin et al., 2013). It is argued that school climate research is rooted in the organisational climate theory and given that the school is an organisation (Tagiuri, 1968). This viewpoint claimed that all school's members experience a similar climate through their shared interaction within the same setting (Van Horn, 2003). On the other hand, other researchers claimed that school climate is an individual-level variable (James, 1982; Miller & Fredericks, 1990). This perspective indicates that each member of the school perceives a different climate based on their background and experience.

These different views have an impact on the assessment and modelling of the school climate (Brault et al., 2014; Van Horn, 2003), since the definition and methods vary and inform different processes (Marsh et al., 2012; Morin et al., 2013). When the school climate is viewed as a school-level construct, it is usually assessed as an average of each perception of school

members within the school, and then the variances between its members are seen as random error. Contrary, when it is viewed as an individual-level construct, the perception of each member is assessed separately, not as an aggregated perception (Miller & Fredericks, 1990). So, each member in the same school may have different values of their school climate that may affect their behaviour in a variety way.

Recently, as some studies have begun to use a multi-level analysis, substantial differences in school climate have been found at both the individual and school levels (Konold et al., 2014). This method of analysis is necessary to precisely investigate the proper unit of analysis for the school climate as it can simultaneously analyse the variance in individual-and school/classroom-level.

3.3.6 School climate factors

In this study, school climate framework that will inform the research design is based on a recent research review by Thapa et al., (2013). They summarised the factors that formed the school climate from empirical research in this field. The empirical review is based on the previous school climate review by Cohen et al. (2009). This theoretical framework is crucial because it is based on the most recent empirical review on school climate and it is comprehensive in the sense that it includes almost all of school climate factors referred to in previous research.

Similarly, as stated before that the factors of school climate that included in this review are relatively agreed, also, since this research uses TIMSS database, the measurement of those aspects is relatively available in this database. However, it is crucial to have a look at how other researchers define school climate domains.

As mentioned before that the school climate could not easily be defined, and therefore, there also no established or fixed set of constructs to clearly identify what is school climate, and what is not school climate (Carrasco Ogaz, 2016; Cohen et al., 2009; Damanik & Aldridge,

2017; Freiberg & Stein, 1999; Payne, 2018; Thapa et al., 2013; Wang & Eccles, 2012; Zullig et al., 2011). The school climate is not defined by a single factor. The juxtaposition of various factors provides support for teaching and learning for all school members (Freiberg, 1998). The Next section discusses the different factors of school climate based on papers cited. The discussion is presented in sequence, based on paper's publication years, and at the end, the school climate factors are compared to look at least the global view and the consistency of school climate factors among the scholars.

Anderson (1982) adopted Tagiuri's (1968) classification of school climate in her empirical review of school climate. There are four keys climate factors described in Tagiuri's model. It reflects the collective attributes of organisational climate, includes ecology (school's physical and material); milieu (the composition of the school's members); social system (relationships amongst school's members); and culture (school's values and belief systems).

Next, Hoy proposed different school climate factor for different school level. So, school climate factors for primary, secondary, and high school are different. As related to this study, for example, there are six factors of climate define the school's health (Hoy et al. 1997) compare to only four factors for high school (Hoy et al., 2002).

That study only presents school climate factors for secondary school. As stated before, there are six factors of school climate in a secondary school which divided into three levels (Hoy & Hannum, 1997). *The technical level includes* academic emphasis and teacher affiliation. (1) The emphasis in academics applies to how much learning quality is guided by the school. It should be high enough to achieve. (2) Teacher affiliation, where teachers feel connected to school, have a strong relationship with other school members, proud to be a teacher and enthusiastic.

At managerial level includes (3) Collegial leadership, where the principal is open, supportive, friendly, and treat others equally and at the same time, the principal sets target and

share the goal to all school members. (4) Resource support refers to the sufficiency of the classroom and teaching materials. (5) Headteacher influence is the ability of headteachers to influence school activities and take actions.

Next, at the institutional level that includes: (6) Institutional integrity refers to the ability of a school to maintains educational integrity. All the factors are combined to measure the overall index of general school health (Hoy & Hannum, 1997).

Freiberg (1998) comes with different term of school climate factors. First, *cohesiveness* refers to the form a positive unity among school members and is committed to improving education. It is much more like an interpersonal relationship, feel like a part of a group. Fraser et al. (1988) stated that a cohesive school climate can be described as how helpful and supportive teachers are to encourage student learning. It is also can be explained by the willingness of school members to work together if needed. Cohesiveness also reflects strong relationship between school members that measures the expected degree of student proximity (Loukas & Murphy, 2007). Loukas (2007) added that the cohesiveness between school members will protect young people from behaviour problems by an enhanced sense of belonging and involvement at school, decrease conflict and peer rejection, could be protective as can encourage youth to deal with their feelings and behaviours rather than to use situational stress factors.

Second, *respect* is a common emotion when a person sees another person as respectable. Respect refers to people in the school who think other individuals can be deemed truthful or fair to do the right thing. It includes building relationships, compassion and respect between people. It is a key value that can affect people in schools who function and learn. According to Dallimore et al. (2004), a climate where the school members respect one another is conducive to school and class involvement. Cohen et al. (2011) added that a respectful school are socially engaged learning environments in which people feel safe, encouraged, engaged and supportive.

All school members can work together to emphasise the importance of establishing respectful behaviours (Meraviglia et al., 2003).

Third, *Control* is when an individual feels able to control and manage a circumstance as well as the sense of order within the school. It also reflects the extent of the policy and pressure used by the school authorities to maintain control over its members (Fraser et al., 1988). This factor is equal to orderly school and much more related to disciplinary in the school.

Forth, *Violence* relates to the use of force to harm a human person, the use of offensive language and threats are also indirect abuse, and the school with a positive climate should hinder this condition. It means that the school is safe both mentally and physically. It seems particularly important to provide a safe environment to enable pupils to take risks, think creatively and objectively and challenge (Davies et al., 2013).

Fifth, *Physical infrastructure* is described as the buildings and premises of the school, the size of the school and the equipment available to the students to create a stimulatory environment. Freiberg and Stein (1999) argued that a school is not biologically an organic entity, but that it has both organisational and cultural attributes of a living organism. The physical structure can have direct influences on staff and learners. Even the wall speaks (Uline, 2008) and certainly contributes to a pleasing working environment (Fraser et al., 1988).

Another researcher, Loukas (2007) simplified the school climate into three factors that comprise physical, social and academic factors. The physical factor includes the visual arrangement of the school building structures and classrooms, the school size and the student-teacher ratios in the classroom, school class organisation, operational efficiencies of the tools and teaching resources, as well as safety and wellbeing.

The social climate factor involves the nature of the interpersonal relationship between all the school members which include teachers and other staff, student, parent, and community; the level of competition and social connection between students; the magnitude of the

participation of students, teachers and other stakeholders; the performance and equality of the inter-personal relations of students, teachers and other staff (Loukas, 2007).

Academic factor includes the quality of the teaching-learning process; teachers' expectations for students' achievement and monitoring the performance of students and submit reports appropriately to students and parents (Loukas, 2007).

In 2010, Zullig and colleagues (Zullig et al., 2010) conducted an empirical review on commonly climate factors used in research and they proposed five school climates which also reaches an almost comparable summary to previous researchers. The authors contain a collection of constructs as follow: order, safety and discipline; academic outcomes; social relationship within school, school physical facilities; member's school connectedness (Zullig et al., 2010).

Before Zullig et al., (2010), Cohen et al. (2009) classify school climate factors with high focus on the educational policy role of school climate and related study. The safety factor, for instance, is related to school bullying projects, in which the school climate plays a crucial role. Cohen et al.'s (2009) view on safety is also relatively similar to Brookover's (1979) study, with more emphasis on beliefs, attitudes, norms and values of schools (Carrasco-Ogaz, 2016). On the other side, the teaching-learning factors are much closer to the classroom climate research which involves evaluating teaching practices by way of student perceptions, the classroom emotional mood, and the learning outcomes expectations of teachers and students.

Equally, in a broader view, the relationship factor is closer to highlighting the value of social interactions between school members (i.e., Bryk & Schneider, 2004). The interpersonal relationship factor is crucial for school development. It is believed to play a critical part to describe differences between schools, school effectiveness, as well as school improvement. The last factor, the physical environment is far more connected to the ecological perspective,

which answers questions about the relationship between physical resources of student learning results and variables such as school size and design (Cohen et al., 2009; Thapa et al., 2013).

Then Thapa et al. (2013) added the fifth factor, as an addition to Cohen et al.'s (2009) school climate model by conducting a systematic review. Of four factors as Cohen et al. (2009) presented above, Thapa and colleagues added the school improvement process factor. This factor is mainly focusing on how principal improve school process. It is mostly discussed about principal leadership. So, the relationship dimension which is located in teaching-learning factors by Cohen et al. (2009), now has its own factor in Thapa's et al. (2013) concept of school climate.

The most current school climate factors are proposed by Wang and Degol (2015). They conducted an analytical review on school climate research. They collected and reviewed around 327 references and selected the 50 most cited papers as a guide to constructing the factors, and then consulted with school climate research scholars to rectify the classification for reduced selections. In the end, they also suggested four factors which are identical and comparable school climate factors with all the authors mentioned earlier.

School safety relates to physical and social-emotional security that the school requires, as well as the degree of order and discipline of all school members. Community refers to consistency and quality of connections between school members. There are four measurements to the community factor of school climate: interpersonal relationships quality, connectedness, appreciation for diversity, and community collaborations. The academic climate is generally defined by three aspects: leadership teaching and learning, and professional development. This factor is one of the most notable and significant aspects of the school climate relating to forms in which teaching and learning instruction is encouraged in school. Lastly, the environmental surrounding plays a critical role to shape school member's experiences. It relates to the

suitability of the school physical environment, building management and facilities, and the availability and distribution of educational resources (Wang & Degol, 2015).

School climate also become an attention to International large-scale studies in education (e.g., TIMSS and PISA). All the studies emphasise the popular term of 'school climate has been acknowledged as among the most significant features on student's learning outcomes (Mullis, Martin, Foy, & Arora, 2012). Still, all the studies reflect the state of a variety of school climate definitions as previously reviewed. The difference in defining school climate reflects the different stresses on what considers as school climate (Carrasco Ogaz, 2016).

As this study utilises TIMSS 2011, TIMSS's framework of school climate is also reviewed. In TIMSS 2011, school climate was measured using different factor scale (Mullis et al., 2012). (1) Emphasis on Academic Success that was collected from teachers and headteachers. (2) Headteachers spend time on leadership practices was collected only from the headteacher. (3) Safe and Orderly School and the data was collected from the teacher (4) School Discipline and Safety. This scale assessed through headteacher report on the severity of the various disciplines and issues in school safety. (5) Students Bullied at School. This scale was measured by asking the student that based on how often students experienced bullying behaviours in the school. Next, a brief explanation of each factor is explained.

Emphasis on Academic Success

TIMSS 2011 believe that academic optimism can even overcome a student's economic and family background in explaining academic performance. School Emphasis on Academic Success scale has five characterises (Mullis, Martin, et al., 2012, p. 248):

1. Teachers' understanding of the school's curricular goals;
2. Teachers' degree of success in implementing the school's curriculum;
3. Teachers' expectations for student achievement;
4. Parental support for student achievement; and

5. Students' desire to do well in school.

This set of questions were given to headteachers and teachers.

Leadership Activities

In their report, TIMSS also placed leadership as part of the school climate (Mullis et al., 2012) and measured by the scale of time headteachers spent on leadership practices, indicating that school leadership is an essential part of the school climate in TIMSS.

Safe and Orderly School

To collect information on this factor, TIMSS 2011 developed a scale that measures the Safe and Orderly School, and the data was collected from teachers. The questions asked the degree to which they agreed or disagreed with five statements, for example, "I feel safe at this school"; "The students behave in an orderly manner".

School Discipline and Safety

This factor measure principals' perception to the degree to which a series of discipline, disorderly, and bullying behaviours are problems in their schools, including arriving late at school, absenteeism, classroom disturbance, cheating, profanity, vandalism, and so forth.

Students Bullied at School.

This factor was measured by asking the student that based on how often students experienced bullying behaviours in the school. The data were collected by asking the student how often they experienced bullying behaviours. For example: "I was made fun of or called names"; "Someone spread lies about me".

There are several previous research studies which reported using TIMSS database and school climate as explanatory variables to explain students' achievement (Mohammadpour, 2013; Wang et al., 2012). Mohammadpour (2013) for example, explored the variation of science achievement of Singaporean students as a function of student, classroom, and school-level factors using the TIMSS 2007. One of the factors assessed was school climate which was measured at the teacher level and school level. However, the author did not use any theoretical

framework of school climate research, and it is only based on what has been measured in TIMSS. Contrary, this research will put the measurement of school climate into different school climate research conceptual framework.

3.3.7 Putting it all together

Table 3-5 shows the comparison of all the school climate factors propose or utilised by the previous researchers and Thapa et al. (2013) as the anchor. The table describes the variability of the terms used by authors, but almost all of them have similarities in what is seen as critical factors. The relationship factor is the most consistent factor that emerged from all the authors. The term used is slightly different, but the terms can be referred as the relationship factor. Factors like safety, teaching and learning, institutional environment are varying among the authors. The least agreed factor is school improvement process, and there is one other author mention this factor.

Those different school climate factors reflect the multidimensional concept of school climate. However, the different categorisations seem to overlap with each other and incorporate intersecting constructs within them. Variations in the school climate are not defined by a distinctive concept in a factor. The categorisations are analogous to permeable categories, which link different features with unequal similarities.

Table 3-1: School climate factors comparison

Thapa et al. (2013)	Safety	Teaching & learning	Relationship	Institutional environment	School improvement process
Anderson (1982)	social system; milieu	-	ecology; culture	-	
Hoy et al. (1997)	-	academic emphasis; institutional integrity	teacher affiliation; collegial leadership	resource support	principal influence
Freiberg (1998)	control; violence	-	cohesiveness; respect; trust	physical infrastructure	-
Loukas (2007)	physical	Academic	social dimension	physical, include safety	-
Cohen et al. (2009)	safety (physical and emotional)	teaching and learning	relationship	environment (physical)	-
Zullig et al. (2010)	order, safety, & discipline	academic outcomes	social relationships, school connectedness	school facilities	-
Wang and Degol (2015)	safety	academic	community	institutional environment	-
TIMSS 2011	Safe and Orderly School; School Discipline and Safety; Students Bullied at School	Emphasis on Academic Success	-	-	Leadership Activities

3.4 Learning outcomes

Student achievement is the centre of almost all aspects of education that influences the effect of all educational improvement efforts (Bloom, 1976; Gardner 2006; Guskey, 2012). However, achievement term is still considered inconsistently between the policymakers and researchers, and arguably, there is no common understanding of what it means (Guskey, 2012). By tracing the meaning of the word 'achievement', Guskey (2012) supposed that it means 'the accomplishment of something'. In an educational context, he argued that 'something' generally refers to articulate learning goals. Therefore, he concluded that 'student achievement is a multifaceted construct that can address different domains of learning' (p.5).

However, the emerging body of SER mainly focuses on the cognitive outcomes of students (i.e., academic achievement) as children go into school to study something that they could not achieve somewhere else (Creemers & Reezigt, 1999). It is argued that academic outcomes are the strongest way to reflect the social duty of schools and the area where the school can clearly distinguish (Opdenakker & Van Damme, 2000). In line with researchers in the SER field, Gardner (2006), who introduced the concept of multiple intelligences, reinforced this view. He reasoned that the fundamental purpose of education is to develop students' way of thinking and habit of thought of the academic subject. If '*it has failed, it is failed in a fundamental way'* (p. 5). From this point of view, it seems reasonable to argue that SER is mainly interested in using cognitive outcomes to assess school effectiveness. In addition, there are also no straight standards and the insufficiency of valid and reliable instruments for measuring non-cognitive outcomes (Hattie, 2012; Luyten et al., 2005).

Though this is a simplistic interpretation of the literature, many researchers have argued that giving attention exclusively on academic achievement would lead to less support of the student non-academic outcomes (Creemers & Kyriakides, 2010; Knuver & Brandsma, 1993; Mortimore, 1988; Thomas et al., 2000). Guskey (2012) claimed that student achievement is a

multi-dimensional concept that can span various fields of learning. In a longitudinal study, Mortimore (1988), from 220 teachers at 50 London primary schools that have been interviewed, more than 77% of the teachers had the intention to develop social-emotional skills of their students. Of the example, in any assessment of the SER, the affective domain is of critical interest in learning purposes.

Indeed, recent academic review of educational effectiveness research stated that SER should give more emphasis on social and affective outcomes (Reynolds et al., 2014). Cohen (2006) proposed that learning goals are not just academic but also social and psychological. Initially, Bloom (1976) suggested that the learning domains are not only cognitive, but it also includes affective and psychomotor. Affective relates to students' views, attitudes and behaviours regarding a specific topic, as well as all processes in the classroom. The self-confidence, self-concept, motivation and other variables of student personal growth may assess these psychological effects of learning (Guskey, 2012). Such learning outcomes are not only significant in their own right but also expected to encourage academic achievement. (Baker et al., 2013; Elias, 2003; Watson et al., 2012).

Similarly, Cohen (2006) stated that the learning objectives should be social-psychological as well as intellectual. Recently, Reynolds et al. (2014) emphasised that the purposes of education should pay more attention to social and affective outcomes. Also, UNESCO (2014) underlines the importance of non-academic learning outcomes as one of the *Post-2015 Education Indicators*. This study, therefore, plans to examine more extensive students' learning outcomes as dependent variables.

Armstrong (2006) assumed that teachers and schools are becoming more interested in academic growth rather than child self-development that he called as the Academic Development Discourses (ADD). He believed that this discourse separates the entire child's curriculum from a solely academic viewpoint and does not integrate with the needs of child

development. Accordingly, standardised test outcomes, and adequate academic achievement progress are the key interests of this discourse. In contrast, Human Development Discourses (HDD) strive to help, promote, and enable the development of an individual as a whole person, including of intellectual, emotive, social, moral, innovative and creative as well as spiritual development (Armstrong, 2006; Indonesia, 2003; UNESCO, 2004).

3.4.1 The non-cognitive aspect of learning

As stated before, student's achievement is a multidimensional construct and can refer to the various domain of learning outcomes (Guskey, 2012). It can take the form as academic and non-academic outcomes. SER researchers also consider this type of learning outcome in their research, i.e., Reynolds et al. (2014) emphasised social and affective outcomes in their recent empirical review of SER.

Affective outcomes refer to students' feeling, attitudes and beliefs on a specific subject, as well as entire progress that they have challenged in the school (Gutman & Schoon, 2013). Moreover, Gutman and Schoon (2013) defined the non-cognitive term refers to 'a set of attitudes, behaviours, and strategies that are thought to underpin success in school and at work, such as motivation, perseverance, and self-control' (p. 2). Inline, Guskey (2012) stated that affective outcomes of learning take forms like student self-confidence, self-concept, motivation, and other aspects of students' personal development. These affective outcomes also encourage academic achievement positively (Baker et al., 2013; Elias, 2003; Watson et al., 2012). Gutman and Schoon (2013) reviewed experimental and quasi-experimental studies that related to the term 'non-cognitive skill'. They examined eight aspects of non-cognitive skills which have identified as a potential key for young people. These aspects include self-perceptions, motivation, perseverance, self-concept, metacognitive strategies, social competencies, resilience and coping, and creativity.

Of the eight aspects, self-perceptions - that including self-concept of ability and self-efficacy had more empirically sound concept compare to other skills. It is because these two aspects of self-beliefs have several advantages as follow:

1. High quality of measurement (indicated by widely used validated measure);
2. a medium strength of evidence as seen on several large-scale meta-analyses of experimental studies;
3. a high (for self-efficacy) and a medium (for self-concept) malleability as determined by the average effect size of its improvement in experimental studies;
4. had effects on other outcomes as they have a medium effect size.

For that reason, this research will use these non-cognitive outcomes of learning called self-beliefs. Also, because this study utilised TIMSS 2011 data, these outcomes are already collected in the student background questionnaire. As an indication of the emerging policy attention in a more comprehensive meaning of learning outcomes, large-scale international assessments like TIMSS and PISA have established broadening their focus to measure not only academic achievement but also in the non-cognitive outcome (Bertling et al., 2016; Bong & Skaalvik, 2003).

Next section discusses each feature of the non-cognitive/affective in term of academic self-beliefs and its derivatives, self-concept and self-efficacy, in detail.

3.4.2 Academic self-beliefs

As stated earlier, self-concept and self-efficacy are two concepts of self-beliefs that attracted a lot of researcher attention (i.e., Bong & Clark, 1999; Bong & Skaalvik, 2003; Gill et al., 2008; Greene et al., 2004; Hughes, 2012; Jiang et al., 2013; Marsh & Martin, 2011; Patrick et al., 2007; Seaton et al., 2014; Valentine et al., 2004). The impact of each self-belief has been provided by those previous studies. However, Bong and Skaalvik (2003) struggles to untangle the separating characteristics of the two belief systems. It is not easy to distinguish

explicitly and conclusively between the two concepts between self-concept and self-efficacy. Nevertheless, some similarities and variations between these two concepts can be illuminated.

3.4.2.1 Self-concept

Self-concept is defined as a self-belief of individual ability formed through interactions with the environment, including school (Burns, 1982; Hughes, 2012; Lepola, 2000; O'Mara et al., 2006; Samuels, 1977). Formerly, Coopersmith (1967) proposed that self-concept was a unidimensional construct. He argued that the different features of self-concept are strongly dominated by a general element that could not be distinguished. Conversely, the unidimensionality of self-concept has been challenged by some researchers (Byrne, 1984; Marsh, 1990b; Marsh & Shavelson, 1985; Shavelson et al., 1976). They argued that self-concept is a multidimensional construct. Generally, it is assumed that specific domain self-concept (e.g., academic, social, emotional, physical,) are structured in a hierarchical arrangement with the general self-concept at the top of the hierarchy (Marsh, 1990b; Shavelson et al., 1976). In this research, academic self-concept will be used, and it is defined as students' knowledge and perceptions about themselves in educational environments (Byrne, 1984).

There is plentiful evidence showing that academic self-concept is an important outcome variable (Marsh, 1990a; Marsh & Martin, 2011; O'Mara et al., 2006; Parker et al., 2014). In some cases, self-concept is even more central than standardised achievement in predicting long-term academic performance (Parker et al., 2014).

3.4.2.2 Self-efficacy

On the other hand, self-efficacy represents the capabilities of a person to judge their abilities in organising and executing courses of action necessary to attain certain types of action (Bandura, 1997; Bong & Clark, 1999). In particular, academic self-efficacy refers to students 'belief that they can effectively complete specific academic tasks (Pajares & Urdan, 2002; Schunk, 1991, 2001). Self-efficacy is different conceptually and psychometrically from

associated motivational concepts, for instance, self-concept or locus of control, because self-efficacy is an achievement-based measure of perceived competence (Schunk, 1991; Zimmerman, 2000). Contrarily, self-concept is a global measure of self-belief, and are not always associated with school success (Marsh, 1990a; Marsh & Martin, 2011; Zimmerman, 2000).

To sum up, academic self-concept and academic self-efficacy are two essential aspects of self-belief in the educational setting. They are explicitly formed toward academic domains (Bong & Skaalvik, 2003; Marsh, 1990b). Concerning the significance of these two concepts of self-beliefs, Parker et al. (2014), by using Australian PISA 2003 dataset, found that these two types of self-beliefs have strong relationships with academic achievement. Both academic self-concept and academic self-efficacy are independent and equally robust as the predictors of tertiary entrance ranks at the end of high school. Self-efficacy is an essential predictor of further education, while self-concept is a significant predictor in choosing a subject in higher education.

3.4.3 Relationship between achievement and academic self-concept

Previous research supports a positive correlation between performance and academic self-concept (e.g. Huang, 2011; Marsh, 1990a; Trautwein et al., 2006). This research has found similar findings under a variety of contexts. For example, studies conducted in various countries, i.e., Arab countries (Abu-Hilal et al., 2013), China (Chen et al., 2013; Yeung & Lee, 1999), and Germany (Trautwein et al., 2006) have demonstrated that students with higher abilities have higher academic self-concept. Those studies also showed the positive correlation between academic success and academic self-concept, providing evidence of the influential research base in this area.

Byrne (1984) argued that academic self-concept has motivational properties and therefore, if academic self-concept changes, it will lead to changes in academic achievement.

In contrast, academic self-concept, according to Calsyn and Kenny (1977), arises primarily as a consequence of academic achievement. Other studies also support the causal relationship between this self-belief and academic achievement (Byrne, 1984; Calsyn & Kenny, 1977; Marsh, 1990a; Marsh & Martin, 2011)

Because of many previous pieces of research support the causal relationship between achievement and self-concept, this research, therefore, treats self-concept not only as an explanatory variable but also as the outcome of learning.

3.4.4 The importance of school as the social context in self-beliefs development

Environmental supports as well as significant others' role such as parents, peers and teachers are essential in forming and shaping self-concept (Shavelson et al., 1976) In a long period, social interactions and an evaluation of person itself, specify the person's views of their performances, successes and failures, leading to an internalised self-representation. Though, from a situational viewpoint, in some situations, explicit feedback and environmental factors could be of more effect than others. For example, in the school setting, the impact of teachers and peers may be of more significance to self-concept than parents (Meeus, et al., 2002).

Since the lives of most children are largely surrounded by home and school environments, it seems natural that school would be significance and salience in developing self-concept. Additionally, development through adolescence suggests a transitional phase when persons are proposed to experience more significant separation from parents (Côté, 2009; Grotevant, 1998; Meeus et al., 2002), leading to school social setting gaining more psychologically significance. Along with this, teachers and their teaching/instructional methods also play as an important social context of student self-development within the school. The teacher also argued to have more effect to student learning and their self-development than any other school factors (Anderson et al., 2004; Cheng, 1994; Creemers & Reezigt, 1999; Marks, 2000).

3.4.5 Learning outcomes measure in international study and some critics

Since this study utilises TIMSS database, it is also important to briefly review and critique the way learning outcomes are assessed in the international large-scale assessment in education (ILSA). In the last several decades, numerous ILSA studies have been conducted (e.g., TIMSS, PIRLS, and PISA), to first investigate disparities in mathematical skills and afterwards investigate measures among students from different countries to consider the impact of variables, such as teachers' instructional strategies, home and school resources and school environment structure and management (Cai et al, 2017). Among the influential studies are TIMSS and PISA.

TIMSS is a cross-sectional international assessment carried out by the International Association for the Evaluation of Educational Achievement (IEA) every four years, conducted for the first time in 1995 (Martin & Mullis, 2013). PISA (Programme for International Student Assessment) is conducted by OECD (Organisation for Economic Co-operation and Development) every three years. PISA is intended to evaluate educational systems by assessing performance on mathematics, science, and reading of 15-year-old student (OECD, 2009).

TIMSS aims to acquire in-depth explanation of the effects of policies and practices within and between educational systems by delivering large-scale comparative studies of schooling outcomes and their other aspects (Foy et al., 2013). TIMSS provides the database for researchers and analysts to examine student achievement of 63 countries participating in TIMSS 2011 (Mullis et al., 2012), including Indonesia. TIMSS also do the follow-up of its study with Video Study since 1999. It aims to compare maths teaching practices in some countries like the U.S, Japan, the Netherlands, Switzerland, and Hong Kong, as well as Indonesia (Hiebert & Stigler, 2000; Ragatz et al., 2015).

Not only measured academic achievement (Mathematics and Science), through the use of questionnaires, TIMSS also assessed student confidence, value, likeliness on a particular

subject as well as other teaching-learning backgrounds. The questionnaires are taken from students, teachers, and headteachers, also country study coordinators which presented with questionnaires that included comprehensive information about the scope and overview of the curricula designed and applied within the education system (Mullis et al., 2012). There are many studies that use ILSA background questionnaire to get information of non-cognitive outcomes like student's self-concept and self-efficacy of education as well as its relationship with academic outcomes (Abu-Hilal et al., 2013; Mohammadpour, 2013; Parker et al., 2014; Wang et al., 2012)

Evidence of ILSA as in TIMSS and PISA has been used to inform education policy in several countries, among the countries are Australia, Japan, New Zealand, India, Singapore, and Indonesia (Best et al., 2013; Tobin et al., 2015)

Tobin et al. (2015) compiled studies in how some countries use the report. For examples: to monitor the quality of their education system (Japan); monitoring achievement differences among student from different SES background (Australia); to provide support in prioritising particular education reform (South Korea); develop country-owned performance test items (Iran); promote better performance standards of students (Russia); Enhancing science learning activities by increasing the frequency of using experimentation and computer software (Malaysia). In the Indonesia context, the use of ILSA to inform policy is not clear, there is no formal documents or reports that aimed to observe the results in-depth. However, when introduced the implementation of new 2013 its national curriculum, the Ministry of Education and Culture stated that TIMSS and PISA results were mentioned as the reason for curriculum change (Mustafa, 2014; Rudhito & Prasetyo, 2016)

3.4.5.1 The limitation of International large-scale assessment of learning outcomes

In spite of the usefulness of ILSA in informing educational policy, there are some controversies and debates on those studies. In Turkey for example, (Gür et al., 2012) pointed

out the misuse of PISA by the government. The government reacted to the PISA result by the intention to change the curriculum, but Gür et al. (2012) argued that the national curriculum should not be used only to meet PISA criteria without further analysis. The condition quite the same in Indonesia context as mentioned above.

Next, the more technical critics also appear, where a variety of challenges have been found that make it difficult to globally compare survey results, which include content localisation or translation, the measurement model, sampling and representation, domain of testing, and validity (Baird et al., 2011). Though the limitation presenting below is not specific on ILSA study, by presenting the limitations, the researcher should be cautious when presenting results from this study (Goldstein & Thomas, 2008). Some limitations are described below:

3.4.5.1.1 May not reflect the curriculum

TIMSS is indeed designed to closely associate with mathematics and science in participating education systems and is intended to reflect the school-based learning of students (Stephens et al., 2016). Some studies found that there is a positive correlation between international large-scale assessment and national assessment scores, as in Canada and Sweden (Cartwright et al., 2003; Wiberg, 2019). However, because it is a cross-countries study, TIMSS delivers similar test across countries, and therefore, the TIMSS result may not reflect what the student has learned, particularly in Indonesia. As Goldstein (2004) argued, the selected items (from items bank) might be selected and administered, then given to student who has not been taught about the tested material. As a consequence, the students may have achieved lower than expected due to the relative severity of the test rather than the failure in the test (Goldstein, 2004).

3.4.5.1.2 The nature of the data

ILSA studies such as TIMSS and PISA only provide cross-sectional data rather than longitudinal. The data provided in the longitudinal studies are distinct from cross-sectional ones. The main difference is that cross-sectional studies only collect data once, while

longitudinal studies observe the same subject group over time (Payne & Payne, 2004). Therefore, the evidence that collected from cross-sectional studies cannot inform the causal relationship between the investigated variables and should be interpreted cautiously. The absence of longitudinal data and potential poor correlations between ILSA studies results and country-specific assessments is a very problematic representation of ILSA results (Goldstein & Thomas, 2008).

3.4.5.1.3 Translation issue/cultural specificity

The language used in the assessment materials is primarily English, and the materials are translated for questionnaire use into more than 40 languages. Not only the standard of interpretation, but the underlying features of each different language themselves may pose problems of comparability (Baird et al., 2011; Goldstein & Thomas, 2008)

3.4.5.1.4 The validity

Another problem is the accuracy of the Rasch design. The items may be pretested (although in all countries, since certain countries choose not to bear the costs involved in this phase), one of the aims of which would be to define "differential functioning" in relation with template requirements. Such items would be removed from the questionnaire collection and updated with possible implications for the reliability of the literacy area to be eventually evaluated. More challenges include the degree to which all student subgroups in each country (e.g., boy-girl, ethnic groups, school types) can still be "model suited" (Baird et al., 2011)

3.4.5.1.5 Sampling

In certain countries the target population for ILSA studies may be far smaller than the current population, and because students who are not in school are typically from low socioeconomic backgrounds, the outcomes of population achievement given by PISA, for example, are not directly comparable to those of countries covered by approximately 100% (Best et al., 2013). Inequalities also happen occasionally if students refuse to participate in the assessment or pupils are missing on the day of the test. Statistically, these are resolved through

weighting of data. However, no quantity of information weighting can monitor "invisible" anomalies in a study, as are the best achieving absent students (Baird et al., 2011).

3.4.5.1.6 Motivation in doing the test

The motivation of doing the testing of students is another yet important issue and the assessment format used in the surveys is widely known, as well as a relatively small number of items (Baird et al., 2011). The unknown degree of motivation for the testing may pose a potential threat to the accuracy of analysis and use of evaluation findings (Goldstein & Thomas, 2008).

3.4.6 Choosing TIMSS over PISA

TIMSS and PISA studies are designed somewhat differently. The main differences are TIMSS samples entire classes and conjoin students to the teacher or classroom level. On the other hand, pupils that participated in PISA are spread across classrooms. Therefore, in PISA, the pupil cannot be grouped to their class or teacher (Swensson, 2017).

Additionally, TIMSS gathers questionnaire data from students and relate it with their teachers. Contrary, PISA has not collected data at the teacher/classroom level. PISA samples are individual students, not classes, and only gathers questionnaire data from students. Therefore, TIMSS data is more suitable in this research to answer its research questions.

3.5 Chapter summary

This chapter explains and justifies the selection of the conceptual and theoretical framework that inform the research questions and design, as well as the methods of data collection and analysis. The selected literature review explained in this chapter are aimed to provide a good understanding of the concept and framework used in this study to answer research questions. Overall, this study recognises the value of the dynamic model of educational effectiveness as proposed by Creemers and Kyriakides (2008), in emphasising the dynamic relationship between factors in a school, which are multilevel in nature. Moreover,

for the current study the dynamic model effectiveness evaluation dimensions are seen as particularly useful and relevant in analysing the nature and extent of school climate factors. To overcome the lack of theoretically driven models in measuring school climate as occurred in most of SER studies, this research utilises school climate factors as reviewed by Thapa (et al., 2013). This school climate framework covers all school climate factors in the previous studies and also most of its consistent factors.

In term of measuring achievement, this study looks at student achievement in a holistic view (Bloom, 1976; Guskey, 2012) which includes academic achievement and self-beliefs. Moreover, for the self-belief outcome, they are seen not only as learning outcomes but also can influence academic performance. In this respect, the study draws the following research questions that need to be resolved in the context of Indonesia based on the reviewed literature.

To build on previous research and address the gap in the literature review and specifically a lack of relevant research in the Indonesian context, the study aims to explore school climate factors and student learning outcomes relationship using the combination of two theoretical frameworks. DMEE (Creemers & Kyriakides, 2008) evaluation dimensions (frequency, focus, quality, stage, and differentiation) is utilised to evaluate the school climate factors as reviewed by Thapa et al. (2013).

The study aim is addressed the following specific research questions:

RQ1: What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?

RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)?

Chapter 4: Methodology

4.1 Introduction to the chapter

The researcher has presented an overall introduction, rationale and overview of the present thesis in Chapter 1. The researcher then described the context of the present study in Chapter 2. Subsequently, the researcher has presented his critique of past research in Chapter 3. In doing so, the researcher has identified the gaps in past research, which led the researcher to propose the following research questions:

RQ1: What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?

- RQ1.1: What is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?
- RQ1.2: After controlling for student, teacher, and school characteristics, what is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?
- RQ1.3: What are the school climate factors that significantly explain the variance between school and classroom performance among Indonesian Year 8 students in math and self-beliefs before and after adjusting the characteristics of the student, teacher, and school?
- RQ1.4: Why do students from the general school have higher achievement and self-beliefs than those from *madrasah* before and after controlling school climate and other factors?

RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)?

- RQ2.1: (a) What are the similarities and differences in school climate between high and low-performing schools? (b) and between religious/non-religious school settings?
- RQ2.2: Are there new factors that can be obtained from a qualitative inquiry that are relevant to highlight differences between the high and low performing schools?

A clear and justifiable research design is required to answer those RQs. But as explained by Morgan (2014), designing and planning research is not a straightforward task. The author had to consider the nature of these research questions, the procedure and means of data collection and analysis, as well as the way the results would be interpreted. Moreover, Creswell (2009) explained that every researcher has their unique worldview and philosophical assumptions that would affect all their research approach. In this present chapter, the author describes the justification of the philosophical, methodological, means of data collection and analysis, as well as some ethical issues arising from the chosen research approach.

4.2 The philosophical approach to research design

4.2.1 Paradigm polarity: Positivism vs interpretivism

In the history of science, how knowledge is constructed has been debated. The debate mainly involved two key schools of thoughts: (1) positivism and (2) interpretivism (Guba & Lincoln, 1994). These two oppositions have promoted the 'incompatibility' or 'incommensurability' mindset among researchers (Howe, 1988; Oberheim, 2016). For the incompatibilist, researchers must constrain themselves to a specific paradigm. They could either be a positivist who uses a quantitative research approach, or a constructivist who uses a qualitative research approach (Howe, 1988). This mindset of incompatibility has led to a fragmentation and polarity in Social Science (Johnson & Onwuegbuzie, 2004).

The positivists follow objective ontological and epistemological viewpoints. They believe in the presence of a true reality, which can be measured and accepted without personal bias. The positivists believe that scientific findings from a given context apply to different contexts, regardless of the differences in social, cultural, and historical realities across contexts (Lincoln et al., 2011). In contrast, the interpretivists follow a subjective ontological and epistemological viewpoint. They insist that cultural, historical, and values in a given context matters. They deny the existence of objective truth (Lincoln et al., 2011).

In terms of epistemology, positivists view factual, tightly controlled, and measurable empirical data as a legitimate source of information. For example, Mason (1973) used a laboratory setting to examine teachers' expectations of their pupils. Mason used a video recording about a student engaging in an exam. The participants (teachers) expectation of the observed student was assessed subsequently. Although the laboratory settings had allowed Mason to control much of potential extraneous variables (e.g., noises, the presence of other people, etc.), such an experimental setting was artificial. The student in the video was not a real student and the context faced by the teachers was not a real classroom. Such an artificial setting may differ significantly to a real-life context. Although such a method may facilitate significant information on particular research issues, the context of the data was not directly evident.

In contrast, interpretivism resists objective claims of truth and uses time and context-based knowledge interpretations. They explore the reality in personal narratives, conversations, and experiences of respondents. Expanding on the illustrative example mentioned in the previous, to understand teacher’s perceptions of their pupil achievement in the future, the interpretivists would utilise observation and interview methods. The interpretivists may use quantitative data, but they would interpret such data from their subjective epistemology. This is in contrast to the positivists who would use interview data in an objective epistemological fashion.

While positivists and interpretivists seem to have a dichotomy in epistemology (Lincoln et al., 2011), the pragmatists contend that a spectrum exists between objective and subjective beliefs. Pragmatists opt to discourage the disagreement between the constructivist and positivist. The pragmatists focus on the practical issues concerning human nature, the study problems and the outcomes of the inquiries. They are more heavily critical of the perfect version of the truth (White, 2010). In the whole case, pragmatism suggests the unique combination of subjectivity and objectivity. The main emphasis on transferability offers a paradigm-shift that can evaluate previous theoretical ideas. Such emphasis also enables the establishment of theories based on a specific context that can be generalised to other contexts. For this reason, researchers have argued the days for the incompatibilists is over. Quantitative and qualitative methods could be combined (Tashakkori & Teddlie, 2003). For this reason, the author opted to follow the pragmatic approach, which will be discussed next.

4.2.2 Chosen paradigm: Pragmatism

In this thesis, the researcher decided to follow the pragmatic paradigm. The ontological and epistemological justification for this choice is presented in the following.

Ontology concerns the concept of truth, what is truth, and whether there is objective truth (Creswell, 1994). On another hand, epistemology concerns the assumptions on how the world operates, the commitment held to the specific view, and the way knowledge is acquired (Creswell, 1994; Crotty, 1998; Guba & Lincoln, 1998). In the present research, the author viewed subjectivity and objectivity as inseparable entities. Observations occur in the social context of the observer and those being observed, and people are rarely able to explain their actions or behaviours in detail. The most they can offer is to clarify what they have done by providing an explanation or story (Denzin & Lincoln, 2005). To understand the climate of a school, the author felt the need to consider the perspective of the school stakeholders and the information that quantitative data can offer. This is especially important considering that a

learning process cannot be separated from its context and it occurs through the interactions of individuals and the social context. For formal education, learning mainly takes place in school classrooms. Each school and classroom have a unique climate.

The author also presumes that individual perception of school climate depends on various factors. Studying school climate without its context would not provide a comprehensive understanding of school climate. There are several sources of data, and many people may be involved in providing information about school climate. Each participant who engages in a school is unique and different from one and another. They have their unique ideas, expectations and values. All of these should be taken into account to understand the climate of schools in a given context.

Educational research is always embedded in a social context such as cultural, social, political, and economic contexts (Jones, 1995; Perry & Weinstein, 1998; Roeser et al., 2000). From Dewey's (1902) perspective, experience always involves a process of interpretation. Beliefs must be explained to generate actions, and actions must be explained to generate beliefs. From this ontological perspective, this study was not entirely positivist or constructivist. It lies between the two paradigms – pragmatism, which strongly supports an equal value between the subjective and objective experience when conducting research (Shannon-Baker, 2015).

From an epistemological point of view, the researcher believed that it was more important to emphasise "what works" methodically in the acquisition of knowledge, as suggested by Howe (1988). As such, the current study contradicted the incompatibility perspective. The researcher believed that the construction of knowledge is situational and social and cannot be separated from political and historical influence (Haack, 1976). This point of view is aligned to that of Teddlie and Sammons (2010), who argued that the conflict between quantitative and qualitative approaches is destructive. Both methodologies should be bridged

to create a synergistic understanding (Sammons, 2010). In short, there is a spectrum of objective and subjective points of view in pragmatism. The method to be used depends highly on the research question at hand (Creswell, 2013; Tashakkori & Teddlie, 1998, 2003).

As the main focus of concern, the researcher viewed school climate and its context as dynamic, multi-dimensional. It involves many parties, with social characteristics that vary from one school to another. For this reason, the author considered further use of a mixed-method to answer the proposed research questions. The author believed that the key was to maintain consistency and coherence in using the mix-method as suggested by Holloway and Todres (2003). Specifically, the method chosen must be flexible and sensitive to the context at hand because social phenomena may differ and depend on the availability of resources and the natural complexity of the phenomenon being examined.

Research on school climate that employed quantitative research designs is prevalent and popular (Fraser et al., 1988; Hoy et al., 2002; Opdenakker & Van Damme, 2005; Voight et al., 2013; Wang et al., 2014; White et al., 2014). This method allows for fast and numerous data retrieval. This approach had helped researchers to feel confident about the representativeness of the sample (Silverman, 2013). Considering these advantages, survey methods remain widely used in social research (Robson & McCartan, 2016). For this reason, the author viewed a quantitative approach as appropriate. However, data collected from a quantitative approach are often too general and broad (Silverman, 2013; Yin, 2014). It cannot explain in detail the processes, formation, or how something came to be. Moreover, the complexity of individual experiences or personal views cannot be explored using only questionnaires with a limited number of questions and room to respond (Yin, 2014). For this reason, the researcher also opted for a qualitative approach to yield an in-depth understanding of the problems at hand.

Qualitative research provides an opportunity to examine the phenomena at hand in detail (Ritchie & Spencer, 1994). It is a dynamic approach that enables researchers to better understand new questions (Miles et al., 2014). It also enables researchers to gather detailed information and comprehend the point of view of the informants without imposing the researchers' assumptions (Smith, 1983). Qualitative interviews are flexible to combine perspectives and explain quantitative findings. It allows researchers the ability to provide the chance for respondents to talk, request information, and clarify their opinions (Knapik, 2006; Smith, 1983). Researchers can also alter their questions, rearrange, ask details in different ways, or challenge them to test whether their understanding is accurate.

By considering the drawbacks and benefit of each qualitative and quantitative approach as described above, the author believed it was best to combine both approaches. Specifically, the researcher believed that if he restricted himself exclusively to either positivism or interpretivism, or using only either quantitative or qualitative data collection and interpretation, he would not be able to achieve the goals and objectives of the present work. For this reason, relative to the positivists and interpretivists paradigm, the pragmatic paradigm chosen for this research was justified and favourable for its flexibility.

The use of the pragmatic point of view in this research was also justifiable if compared to the critical realism point of view. Similar to the pragmatic perspective, critical realism is also, in essence, a mixed-method research's philosophical foundation (Maxwell & Mittapalli, 2010; Lipscomb, 2011). This point of view sees reality in terms of the interaction between subjectivity and objectivity, emphasising that people have their unique view of their world (Crotty, 1998). Critical realism is also meant to be both anti-positivism and anti-interpretivism. However, As described by Wilson and Greenhill (2004) the general aims of critical realism entail a commitment to emancipation, a focus on issues of equality and inequality, and

questioning the status quo (Wilson & Greenhill, 2004), which was not the aim of the present research.

4.3 Methodological approach

4.3.1 Mixed-method research (MMR)

Consistent with the pragmatic epistemology, MMR was adopted to answer the research questions of the present study. MMR is an advanced methodological movement and mode of investigation in SER (Sammons, 2010). This technique integrates two or more types of research data collection strategies (Tashakkori & Teddlie, 1998). The most interesting is the way MMR offers various techniques for summarising data by combining and mixing methods as a basis for policymaking and practice (Simons, 2013). The present research used two different methods (quantitative and qualitative) to understand the results more thoroughly. However, Greene and Caracelli (1997) argued that "using various methods at the same time may not guarantee proper research" (p. 107), the author believed that it was better than to rely on a single method.

The advantages and disadvantages of the mixed methods approach have been widely discussed (Creswell, 2003). This method provides details on quantitative findings through triangulation with qualitative evidence. This design can be very helpful when unexpected results emerge from quantitative analysis (Morse, 1991). Additionally, this research was intended to measure school processes (school climate). Researching school processes utilising only quantitative measures (i.e., survey) alone is not enough considering that quantitative data does not provide the best measures of the school process (Teddlie et al., 2000). The qualitative inquiry in the present research was meant to support the interpretation of the quantitative analysis of the present research.

In particular, the present research used a sequential explanatory design of MMR. In the first phase, a quantitative study was conducted, followed by a qualitative study. The assumption

here was that quantitative analysis offers a good general understanding regarding the research questions. The qualitative analysis would then sharpen the findings through further exploration of the opinions of respondents (e.g., Creswell, 2003, 2009, 2014; Tashakkori & Teddlie, 1998).

4.3.1.1 Mixed methods research design: Integration, priority, timing and mixing

To bring a convincing, systematic and high-quality MMR, it is essential to explain MMR features relating to the following four criteria: integration, priority, timing and mixing (Creswell & Plano Clark, 2011). Which will be discussed in the following.

Integration discusses the specific crossing point between quantitative and qualitative methods. In this study, findings from the quantitative phase were not only used to estimate the range, school climate and math student self-beliefs in Indonesia (RQ1), but also to determine schools for the second phase (RQ2). This study linked the analysis of quantitative survey (TIMSS) data to the collection of qualitative data. The analysis in phase one (quantitative) guided the sample selection for the qualitative phase.

Priority relates to the emphasis given to each method. In this study, a noticeable quantitative method was selected when developing a methodological approach to answer RQ1. The quantitative process explored the correlation between the school climate and the students' learning outcome and then was supplemented by a qualitative approach.

Timing refers to the use of methods in chronological order. Sequentially, this study began with a quantitative phase and continued with a qualitative stage. The two data types were analysed sequentially to complement the results from one method to another. It matched the order in which the research questions were stated. This design made it possible to investigate and consider first, the dependency in the structure of Indonesian hierarchical education, then, illuminated the quantitative measures of student achievement with qualitative data. The second phase facilitated explanations and provided further insights and enrichment of the interpretation of the results of the first phase.

Finally, *mixing* focuses on the time in which the methods are combined. In this research, mixing took place during the data collection of the qualitative phase to supplement the quantitative phase.

4.3.1.2 Sequential explanatory mixed methods design

Sequential mixture analysis uses several approaches for sequential data collection and analysis (Tashakkori & Teddlie, 1998). Data analysis always starts before the next phase of data collection (Creswell, 2009, 2014; Tashakkori & Teddlie, 1998; Teddlie & Tashakkori, 2008). Creswell and Plano Clark (2011) define SEMMD as a combination of unusual approaches, a mixture of secondary data sets and a qualitative inquiry. They specifically accept this approach because they realise that ordinary designs are often not enough to describe certain phenomena, such as school climate. For this reason, this design is also called a qualitative supplement approach (Morgan, 2007).

In the first phase, data from TIMSS 2011 were analysed. The goal of this quantitative phase was to develop a general understanding of how school climate may be associated (or not) with students' achievement and their self-beliefs, specifically in the context of Indonesia where little comparable research has been conducted, as described in Chapter 1. In the second phase, four case studies were carried out to better understand the relationship between school climate and students' educational outcome. This research sought to explain why this relationship differs across the different schools (which vary by performance and religious type) by exploring participants' views.

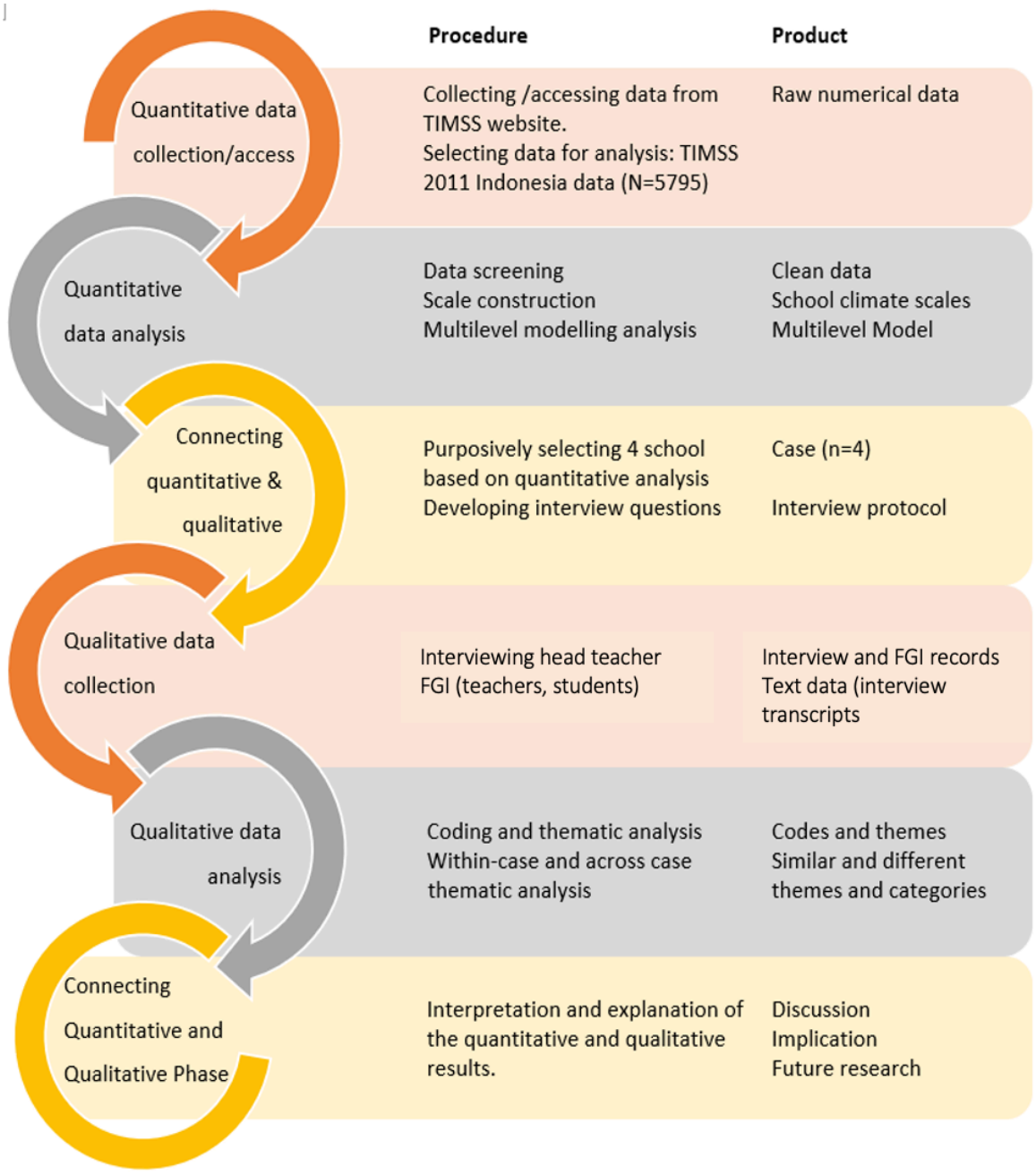


Figure 4-1: The sequential explanatory mixed-method procedure (adopted from Ivankova et al. (2006)

4.4 Quantitative phase

4.4.1 Secondary data analysis

To address RQ1 (What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?) This empirical study analysed secondary data taken from TIMSS 2011. The researcher chose this secondary data approach to gain access to large samples, which ensured the representativeness, generalisation, scope of the study assessment, and to attain an appropriate level of statistical power (Shultz et al., 2005). It is also important to point out that the educational reforms in Indonesia started in the early year of 2000. The TIMSS 2011 data is thus the better choice to be used to examine the effectiveness of the educational reform after about a decade of its establishment. For this reason, TIMSS 2011 data was opted over other data sets such as TIMSS 2015.

TIMSS 2011 data could be considered a form of secondary data. Some central issues when analysing secondary data is that the quality of the quantitative results relies significantly on the accuracy, validity and reliability of the data. But such issues can be dealt through careful examination of the data. Another central issue concerning the use of secondary data is regarding the originality of the research. The author himself was aware that other researchers had used TIMSS 2011 dataset to examine school climate (e.g., Wang et al., 2012; Kyriakides, 2006; Chen et al., 2012; Mohammadpour, 2013; Abu-Hilal et al., 2013). Such research may overlap with the current research. As such, the author felt necessary to ensure the originality of the present research.

Although searching for studies in the Indonesian context that had used TIMSS data set was a challenge, the researcher managed to find several studies that had used this data set. Wijaya (2017) investigated fourth graders' understanding of the basic fraction concepts.

Saputro et al. (2018) examined the rate of errors Indonesian students made when responding to the TIMSS algebra domain problem. Fenanlampir et al. (2019) looked at obstacles hampering Indonesian students to achieve better ranking in TIMSS and PISA. Thomas (2017) explored Indonesian students' performance on TIMSS over time compared to students of other nationalities. After carefully searching for research conducted in the Indonesian context that had used TIMSS data, the author concluded that his research questions were novel and had not been asked by other researchers.

Another limitation of using secondary is that the data may not as ideal as a researcher would want it to be. For instance, the TIMSS data set has only a few items collected for certain school climate factors. This means that the author had to be extra cautious when interpreting the findings. Moreover, the data set was also lacking records about students' past attainment, which could have been used as a substantial indication in the variation of students' learning outcomes. TIMSS also included data from many countries. Due to cultural differences, people from the various parts of the world might have interpreted the questions in TIMSS differently (as mentioned in Chapter 3 about the limitation of ILSA). These limitations had made it imperative for the author to reconstruct the school climate factors provided by TIMSS.

4.4.1.1 Indonesian Sample in TIMSS 2011 data set

TIMSS employed a stratified two-stage cluster sampling approach. Schools were sampled by a systematic probability proportional to size (PPS), then one or two classrooms per school were selected (Joncas & Foy, 2012). In TIMSS 2011, Indonesia participated in assessing 8th-grade students (13-14 years old). 153 schools (out of 36,234) and 5,795 students (out of 3,178,536) participated in the data collection. The sample consisted of 2,823 male students and 2,972 female students. The sample covered schools from 31 of the 33 provinces in Indonesia. The school population was separated into different strata. The stratification process was meant to ensure a representative quantity of schools within each stratum. In the Indonesian context,

the sample was explicitly stratified to public versus private schools and general versus. Islamic schools (see Table 4-1).

It has to be noted that a wide range of quality exists between schools in Indonesia. For this reason, TIMSS implicitly stratified schools not only by province but also by performance. The performance of the schools was classified into three categories: high, medium, and low (Joncas & Foy, 2012). This implicit stratum is nested within the explicit strata (general and *madrasah*) and was used for arranging the sampling frame preceding to the systematic sampling of the schools (Joncas & Foy, 2012).

Table 4-1: Summary statistics of mathematic achievement data

		Explicit Stratum				Total
		Public General	Public Islamic	Private General	Private Islamic	
Implicit Stratum	High	26	3	10	9	48
	Medium	51	4	16	13	84
	Low	11	1	6	3	21
Total		88	8	32	25	153

Source: Methods and Procedures in TIMSS & PIRLS, 2011 (Joncas & Foy, 2012).

4.4.2 Data preparation

The researcher started the analysis of the first phase of the study by cleaning and preparing the TIMSS 2011 data set. Variables were re-coded into new variables and school-level variables were aggregated. In the next section, the researcher explained TIMSS 2011 data treatment

4.4.2.1 Missing data

According to Graham (2008), there are two key methods commonly used to treat missing values: deletion and substitution. The deletion strategies are listwise and pairwise deletion. In the listwise deletion, cases in which information is not present in one or more variables are deleted (Allison, 2001; Dong & Peng, 2013; Pepinsky, 2018). When using pairwise deletion, the analysis is conducted using all cases that have scores for every variable

or combination of variables (Allison, 2001). In the substitution approaches, missing values are imputed. There are two main substitution techniques: single and multiple imputations. A common form of single imputation is by replacing missing values by the average of the score of the given variable (Graham, 2008; Schafer & Graham, 2002; Zhang, 2016). On the other hand, the multiple imputations refer to the process of replacement by a more than two imputed value of each missing value (Allison, 2001; Graham, 2008; Little & Rubin, 2019)

As in a lot of surveys, the TIMSS 2011 data set also contained missing data. In multilevel studies, missing data is a challenging problem because it reduces the number of observations that can be analysed (Leeuw & Meijer, 2008, p. 5). Missing data, especially at the top level can be a problem because any data lost at a higher level (e.g., School) removes all data in the lower unit (e.g., students) from the analysis (Gibson & Olejnik, 2003).

In the present research, mathematics scores were treated as one of the outcome variables at the student level. Fortunately, there were no missing cases on this variable. However, there were missing data on the academic self-concept variable. This might have happened because this variable was constructed from several different items. There were also cases of missing data on the predictors, but the problem was not severe since the missing data were less than 5% (Tabachnick & Fidell, 2007). However, rather than removing observations with missing value, the missing values were imputed by using expected maximisation (Rubin et al., 2007).

Since there was minimum missing data at the student level, the author decided not to impute missing values for the “Student Home Resources” scale. However, as pointed out by Raudenbush & Bryk (2002), the sample size at the school level has more influence on power estimate than the sample size at the student level. As such, missing data at the school level were imputed. Specifically, conditional mean imputation was used to impute missing data for each school level measure with missing data in each country. Following the single non-stochastic regression method described by Little and Rubin (2019), a separate multiple regression

equation was built to predict missing values for one measure conditional on the values of all of the other measures in the model. For example, if a country was missing scores for the "Effective Instructional Strategies for Encouraging Mathematics Reasoning" scale, all of the other school-level measures were entered into a regression equation used to predict the missing score. The missing values for each Rasch scale were then replaced with the predicted values from each model. Although there are more sophisticated methods to impute missing data (e.g., multiple imputations), this straightforward approach was taken due to the minimal amount of missing data.

4.4.2.2 Multicollinearity

Multicollinearity refers to explanatory/predictor variables that correlate with other explanatory variables. Multicollinearity appears when a model consists of multiple factors that are correlated not just on dependent/response variables, but also to each other. Multicollinearity produces shared variance between variables, which reduce the power to predict the dependent variable and the relative functions of each predictor's variable.

There are at least two methods to detect multicollinearity. First by conducting a correlation matrix between predictors variables (Field, 2005); and the second by checking variance inflation factors (VIF). The greater the value of VIF, the higher the correlation of the variable with other variables. However, some scholars differ in a rule of thumb in measuring VIF. For example, Hair et al. (2014) suggested the maximum level of VIF value of 10. Ringle et al. (2015) recommended five as the maximum level of VIF. In this study, multicollinearity was measured by VIF. This study adopted Hair et al. (2014) to check VIF among the predictor's variables.

The author examined the VIF of all predictors. Predictors with VIF over 5 were omitted as they practically measured the same thing. The mean VIF for the school climate factors was

1.24, student's characteristic was 1.12; teacher characteristics was 1.02, and school context variable was 1.14.

4.4.2.3 Data Distribution

In multilevel modelling, an essential assumption of the model is that the residuals are normally distributed (Raudenbush & Bryk, 2002). This study used a Q–Q probability plot to check this assumption by using the residual files generated by MLwiN after the first model was fitted. The normally distributed residuals were expected to lie in a straight 45-degree line. The plot shows that the distribution of outcomes variables was in the straight line and this is the indication that the data distribution was normal.

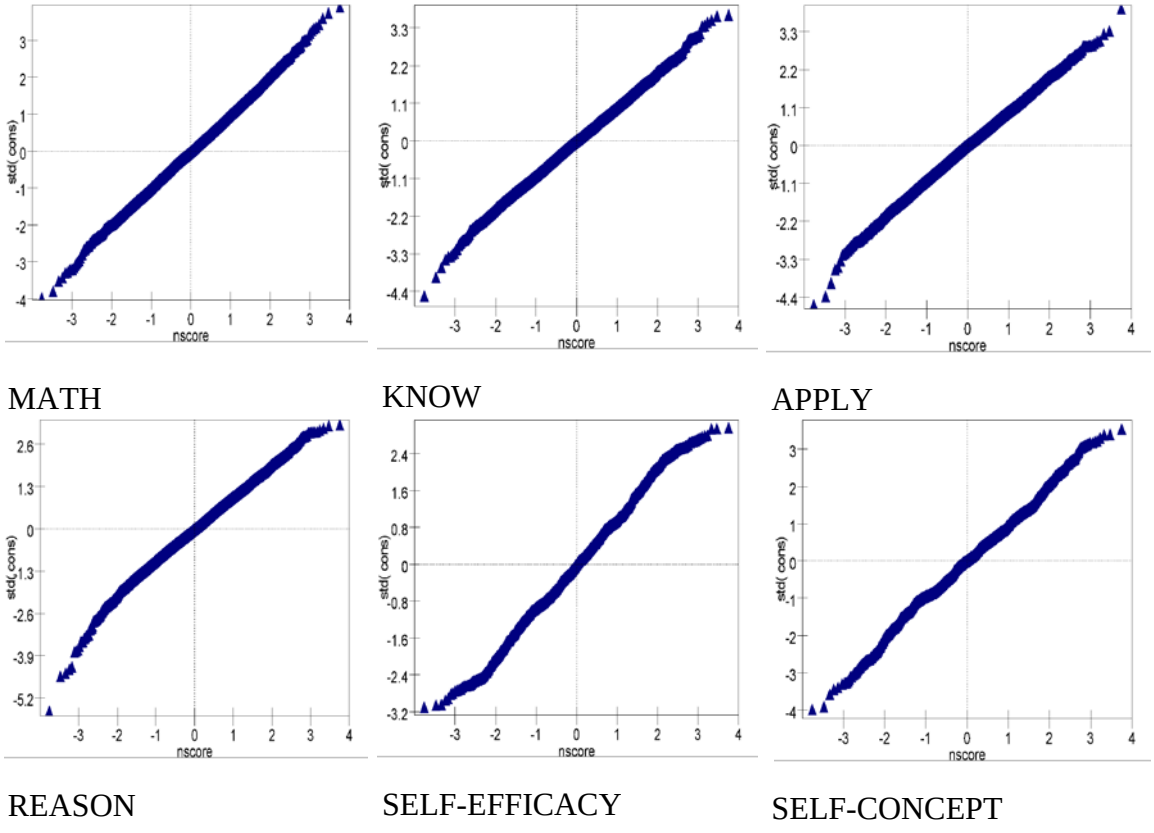


Figure 4-2: Q-Q plot

Note: Based on 1st Plausible value for MATH, KNOW, APPLY, and REASON

4.4.2.4 Sampling weight

One of the essential components for analysing complex sampling data is sampling weights (Asparouhov, 2006; Carle, 2009; Hahs-Vaughn, 2005; Laukaityte & Wiberg, 2018). The weights can be defined as the number of units represented by the respondent in a target population and the purpose is to adjust the sample size, so population estimates can be reduced to accommodate disproportional selection probabilities (Rust, 2014). The parameter estimates or standard errors may be incomplete without taking into account the correct weights (Asparouhov, 2006; Hahs-Vaughn, 2005; Korn & Graubard, 1995; Laukaityte & Wiberg, 2018; Rust, 2014).

In TIMSS study, a multistage stratified cluster design was conducted to enhance the feasibility of data collection. However, the probability for each sample selection would not represent the equal number of students in the population. This means that not all students and schools in a country have the same chance of being included in the study (Joncas & Foy, 2012). Therefore, to account for differential probabilities of selection, and to avoid bias in parameter estimates, TIMSS computed several sampling weights. However, most of the weights (i.e., total student weight) were suitable for single-level analyses but not for multilevel analysis (Rutkowski et al., 2010; Asparouhov, 2006; Laukaityte & Wiberg, 2018). Therefore, in multilevel models, the weights must be scaled differently across clusters so that the sum of the weights would equal some cluster characteristics (Asparouhov, 2006; Laukaityte & Wiberg, 2018; Rutkowski et al., 2010).

Following the suggestions of those researchers, this study recalculated the level 1 (student) weights supplied in TIMSS 2011. The level 1 weight was obtained by multiplying the variables wgtadj2, wgtfac2, wgtadj3, and wgtfac3 (TIMSS provided all these weights), and for level 3 weights SCHWGT does not require recalculation. Moreover, for the classroom

level, the selection was assumed to be equal. Scaling of upper-level weights was not required and should not affect parameter estimates (Laukaityte & Wiberg, 2018).

4.4.2.5 Plausible Values (PV)

TIMSS estimated five plausible values for each student based on their answers to a sub-test of the items. The plausible values were an estimated score of how a student might have achieved if all 190 items were tested. A matrix sampling design was implemented to assemble the items into different booklets because it was not possible to apply all 190 items to the students (Foy et al., 2013). Each student was tested using only one booklet. Each booklet accommodated the plausible items to test a student's plausible achievement score (Foy et al., 2012). Since the uncertainty under these conditions, the distribution of student ability, or its joint distribution with other variables can be approximated using each student's estimated ability (Foy et al., 2012).

Therefore, it is necessary to use all the plausible values when analysing TIMSS data. For example, the simulation study by Laukaityte and Wiberg (2017) showed that some commonly used user approaches give inaccurate results while others give acceptable estimates but incorrect standard errors when PV was not employed. This study thus also used all plausible values (by including each PV in analysis) to estimate math achievement scores, and then combined them by using Rubin's rules.

4.4.2.6 Centring

In multilevel modelling, according to Tabachnick and Fidell (2007), centring is taking away raw scores into its deviance from the mean score. Centring makes the value of scores more interpretable. The interpretation of the intercept is often unreasonable without centring, for instance, if age is a forecaster. Non-centred raw predictors should not be the standard scale, since in the social science attributes usually have no meaningful or real zero (Kreft & Leeuw,

1998; Kreft et al., 1995) The centring decisions should be linked to the research questions or on a theoretical basis (Brincks et al., 2017; Enders & Tofighi, 2007; Kreft et al., 1995).

Since this research mainly aimed at investigating school climate effects, centring was used to simplify the interpretation of the intercepts and separated the within-group effects from the between-group effects (Raudenbush & Bryk, 2002). Accordingly, student-level predictors were class or school-mean centred, and school-level independent variables were grand-mean centred.

It assumed that individual students' learning outcomes are affected by school climate. Student or teacher perception on school climate is regarded as an independent observer of the school, the referent is the school, and responses are aggregated across all students and teachers within a school to provide an indicator of school climate. So, grand-mean centred is ideally suited for this situation (Enders & Tofighi, 2007). For comparison, in the three-level study, Mouhammadpour also used grand mean centred when analysing Singapore TIMSS data (Mohammadpour, 2013).

4.4.3 TIMSS data

Since this study seeks to examine the relationship between school climate and students' achievement as well as their self-beliefs, TIMSS 2011 mathematics achievement and its cognitive domains, student self-concept and self-efficacy (constructed from student background questionnaire) were used in the study. The next section discusses the variable in TIMSS that was used in the study. First, student achievement data, and second TIMSS background questionnaires at the student, teacher, and school levels.

4.4.3.1 Students' mathematics achievement

The TIMSS 2011 8th-grade students' performance consists of mathematics and science. However, for a practical purpose, this study only used mathematics achievement because science class in Indonesian Junior secondary school teaches sciences separately by several

independent lessons such as biology, physics, chemistry, and earth science. Therefore, students' perception of their self-belief also reflects this separation.

4.4.3.2 Mathematics' cognitive domains

The math assessment structure was divided into two sections: content and cognitive. First, the content section relates to the math skills needed to complete the tasks (e.g., multiplication, perimeter) (Mullis et al., 2012). Second, the cognitive section refers to the expected cognitive demand required of a student to respond to a provided mathematics question. The cognitive domain comprises knowing, applying, and reasoning. Those domains classify the cognitive behaviours of students when they react to the content of mathematics. All of these cognitive domains were also included as dependent variables together with math total score to have an idea in which part is Indonesian student better or worse.

The knowing (KNOW) domain is linked to procedures, concepts, and facts that students are supposed to know. It includes sub-skills, such as recalling, recognising, computing, retrieving, measuring and classifying. KNOW comprises 35% of math questions. The applying (APPLY) domain is associated with the capacity of students to use knowledge and maths principles to solve problems. Students are expected to select a correct method, strategy, or math operation, to answer problems. APPLY comprises 40% of math questions. The reasoning (REASON) The domain involves conditions, dynamic structures and multi-stage challenges that are not common, which requires review, generalisation, inference, rationale, and solution to non-routine issues. REASON comprises 25% of math questions.

4.4.3.3 Background questionnaires.

TIMSS pointed the National Research Coordinators (NRC) in each participating country to evaluate the assessment frameworks on which the test items are based. They are also responsible for reviewing the background questionnaires taken by administrators, teachers, and students at participating schools. These questionnaires ask questions about school resources

(e.g., textbooks, technology), time on learning, education of parents, attitudes about the subject matter, frequency of homework, as well as number of books at home (Martin et al., 2012).

Student, teacher, and school questionnaires were given (Mullis & Martin, 2011) to sampled schools. Student questionnaires asked students about their home backgrounds, attitudes toward school, and attitudes and behaviours towards particular subject matter (i.e., mathematics, science, and reading). Teacher questionnaires asked teachers about their feelings regarding the school climate; preparedness to teach mathematics, reading, and science; their education; and their classroom coverage of subject matter tested by TIMSS. Teachers were also given subject-specific questionnaires, where mathematics teachers completed mathematics teacher questionnaires. School questionnaires asked headteachers to answer some questions about student demographics, school resources, educational programs, and school climate.

As the main aim of the research is to examine the relationship between broad learning outcomes and school climate, this research utilised student, teacher, and school questionnaires to construct school climate factor scales. Moreover, the non-cognitive outcomes variable of the study (self-beliefs) were also built from the background questionnaire, particularly student questionnaire. Thus, the next section explains the analysis procedure for re-analysis climate factor scales as well as the self-belief scales.

4.4.4 Re-analysis of climate scales: Rasch Model

Rasch measurement (RM) is a tool for controlling item quality (item fit), construct validation, theory development, and creating standardised measurement (Iramaneerat et al., 2008). Thus, the scale used in this study (school and classroom climate scales and self-beliefs scales) was reconstructed using this approach.

RM uses logit as measurement units, where the logits are obtained by transforming ordinal measure (i.e., the Likert scale) into interval data, then the data is mapped into a linear

scale like a ruler (Bond & Fox, 2007; Iramaneerat et al., 2008). RM evaluations based on a sample of respondents' answer to a collection of measurement scales; the person is then classified by ability, while the items are classified by difficulty (Bond & Fox, 2007; Boone & Noltemeyer, 2017; Ishak et al., 2018).

This research used RM analysis to re-construct and validate scales in TIMSS data before using it into further analysis. As mentioned before, the scale that been validated using RM analysis are self-beliefs scales and school climate and classroom factors scales. For this study, an RM is noticeably appropriate for validating the scales, since TIMSS also created scales using RM, particularly the Rasch partial credit model (Martin et al., 2012). This study, also using the same methods of RM and fitting an item model to a set of data using Winsteps 3.73 (Linacre, 2011b). Winsteps is among the most popular RM software and has been used by many researchers to measure latent variables (e.g., Boone & Noltemeyer, 2017; Cai, 2017; Cavanagh & Waugh, 2011).

Although the use of RM is popular particularly in Australia and the USA, this approach also under critics (Panayides et al., 2010), particularly in the UK. Among the person who mostly has given the critics was Harvey Goldstein (Goldstein, 1979). But the debate still continued, and major international large-scale study uses this approach (Panayides et al., 2010). To get more detail on this debate, please refer to Panayides et al. (2010).

4.4.4.1 Climate scales' goodness of fit

After the RM analysis has run, the validation process will consist of examining output for (1) unidimensionality and (2) item fit to the model. It includes considering the item and person reliability, item infit and outfit statistics (Ishak et al., 2018). First, unidimensionality is generally indicated via *Principal Component Analysis*. Unidimensionality examined whether a group of items measure the same single latent variable (Hattie, 1985).

In this research, *First Contrast Unexplained Variance (FCUV)* was used to detect unidimensionality. The rule of thumb corresponding if the eigenvalue of the first contrast is small, usually less than 2.0 (Linacre, 2011a). The process is repeated until the FCUV of each scale achieved less than 2.0.

Second, item fit to the model, RM analyses generally use two measures of fit: infit and outfit. Table 4-2 shows the general rule of thumb of this measure. Small values indicate that responses are too structured. Large values, generally a more pressing concern, indicate responses are too random (i.e., less discriminating). Infit is an information weighted index, implying it gives more weight to items that are well targeted to a person, whereas outfit is closer to a standard chi-square. In practice, if either of these indexes reveals a problem, an item needs to be reviewed (Linacre, 2011a).

Table 4-2: Explanation of parameter-level mean-square fit measurements

Value	Meaning
>2.0	Misrepresents the measurement system.
1.5 - 2.0	Ineffective for measurement construction, but not degrading.
0.5 - 1.5	Effective for measurement.
<0.5	Less effective for measurement, but not degrading. May create incorrectly valuable reliabilities and separations.

Source: Linacre (2011a)

Additionally, reliability for each scale was also reported. The reliability is divided into two, item and person reliability. For item reliability, low values suggest a narrow range of item measures or a small sample. Generally, low reliability of item, in particular, implies the item stability estimates is low due to small sample size. On the other hand, low person reliability shows limited item numbers. To increase a person's reliability, a person with more extreme abilities (high and low) is required as well as adding more items (Linacre, 2011a).

As stated above, regarding the creation of climate scales, school climate measures have been obtained from the student, teacher, and school background questionnaires of TIMSS 2011. All possible school climate scale collected in TIMSS were reviewed to be included in this

study, after consideration of the overarching school climate factors identified by Thapa et al. (2013), including safety, relationship, teaching and learning process, institutional environment, and school improvement process. However, some school climate factors identified from the literature are not explicitly measured in TIMSS 2011, and some sub-factors, such as student-student relationship in relationship factors cannot be obtained. As a result, not all school climate sub-factors can be analysed, which is a limitation of this study. Next section discusses each of the five Thapa et al. (2013) school climate factors that can arguably be measured via TIMSS 2011 data.

4.4.4.1.1 Safety

Since the TIMSS questionnaire does not separate safety as physical, emotional safety, and rules, this factor will be aggregated as one factor – safety at each level (student, teacher, and school). The measure consists of (a) headteachers' report of discipline problems, (b) teachers' reports of school safety, and (c) students' reports of being bullied at the school. Mostly, the items that measure these factors give more emphasis on emotional safety and how norm and rule are applied. The detail of questions of each report can be seen in Appendix 6.

4.4.4.1.2 Relationship

There are three sub-factors that can be obtained in TIMSS questionnaires to measure this factor, student connected to school (student like school), teacher connected to school and teacher professional relationship/interaction. Next, the relationship factor of school climate is teacher professional relationship. It is measured using one of the teacher questionnaires asking about teacher interaction. This variable is obtained from five items (*e.g., 'work together to try out new ideas'*). It is also scored on a 4-point Likert scale ranging from 1 (never or almost never) to 4 (daily or almost daily).

4.4.4.1.3 Teaching and learning process

The important aspect of this factor is how the school supports students' academic achievement. There are two similar questionnaires conducted, headteacher and teacher

questionnaires which report school emphasis on academic success. This factor is measured by five items (e.g., *'teachers' understanding of curricular goal*'). The responses are ranging from 1 (very high) to 4 (very low). At the student level, the teaching-learning process is measured by student engagement in a math lesson. It *measured using five items (e.g., 'I am interested in what my teacher says'). This factor* is scored on a 4-point Likert scale ranging from 1 (agree a lot) to 4 (disagree a lot).

4.4.4.1.4 Institutional environment

This factor is measured by two questionnaires obtained from teachers' view of their working conditions, and headteachers' perception of the degree to which school facilities and resource availability are affecting the quality of instructions. Teachers' questionnaire contains five items (e.g. *'the school building needs significant repair'*) and reported on four responses ranging from 1 (not a problem) to 4 (serious problem). Next, the headteachers' questionnaire, on the other hand, contains two sections. Six items are asking about general school resources (e.g., *'school buildings and ground'*), and another six items are asking about resources for math instruction (e.g., *'teacher with a specialisation in math'*). The responses are ranging from 1 (not affected at all) to 4 (very affected).

4.4.4.1.5 School improvement process

School improvement process factor is obtained from thirteen items asking about headteacher leadership activities (e.g., *'monitoring teachers' implementation of the school's educational goals in their teaching'*). The responses are ranging from 1 (no time) to 3 (much time). TIMSS has not created this scale, so the researcher created the scale.

Table 4-3 shows the final fit statistics of each factor of the school climate used in this study. Also, all the factor scales are in the acceptable fits statistic as required in RM. After re-analysing the factor scales, the name of factor is modified by looking at the item which constructed the scales.

Table 4-3: Fit statistics for school climate factor scales

Scales	1st contrast loading	Infit (mean)		Outfit (mean)		Reliability		N item
		MNSQ	ZSTD	MNSQ	ZSTD	item	person	
Safety								
Student being bullied in school	1.5	1.04	2.2	0.82	-2.6	1	0.65	4
Teacher feeling safe in school	1.7	0.93	-0.6	0.98	0	0.95	0.49	3
School safety	1.9	1.08	0.4	1	0	0.72	0.85	5
School discipline	1.6	1.08	0.3	1.28	1	0.68	0.84	6
Relationship								
Student like school	1.6	0.99	-0.6	0.89	-3.1	1	0.18	2
Teacher connected with school	1.6	0.97	-0.4	0.97	-0.1	0.97	0.49	4
Teacher interaction	1.5	1.01	0	0.96	-0.4	0.96	0.73	4
Teaching-learning								
Student engagement with math lesson	1.5	1.03	0.4	0.82	-5.4	0.98	0.58	4
Teacher emphasis on academic success	1.7	1.09	0.1	1.03	0.2	0.99	0.75	4
School emphasis on academic success	1.8	1.09	0.8	1.24	1.7	0.97	0.76	5
Teacher engaging	1.6	1.2	0.2	1.12	0.3	0.97	0.13	5
Teacher confidence	1.5	1	-0.1	1.07	0.1	0.93	0	4
Physical environment								
Teacher working condition	1.7	1	0	0.98	-0.2	0.87	0.62	5
School technology support	1.9	1.04	3.4	1.5	3.1	0.86	0.83	4
School general resources	1.6	1.39	2.1	1.16	1	0.93	0.57	5
School improvement process								
School leadership	1.9	1.42	0.9	0.76	1	0.5	0.53	8

4.4.5 Re-analysis of students' self-beliefs scales: Rasch Model

In this research, student self-beliefs were divided into two scales, math self-concept (SELFCONCEPT) and math self-efficacy (SELFEFF). For example, to measure SELFEFF, the students' confidence in mathematics scale was examined throughout student's questionnaire items, for instance, "I usually do well in mathematics". (Foy et al., 2013). These two constructs are measures of different concepts, in brief, SELFCONCEPT emphasise more on the learning outcomes, whereas SELFEFF is more about the prediction of outcome. SELFEFF is utilised to define students' beliefs about whether they can carry out particular math achievement in the future through their efforts, while SELFCONCEPT indicates students' belief in their ability to get certain levels of achievements in math after they evaluate themselves in mathematics (Cai,

2017). For a complete list of questions, please see Appendix 3. The variables were recoded to assist interpretation.

4.4.5.1 Math Self-Concept (SELFCONCEPT)

Math self-concept (SELFCONCEPT) was measured in TIMSS 2011 using six items (e.g., '*Mathematics is harder for me than any other subject'*). In TIMSS, those items are used to measure whether student like learning maths or not. So, it measures student's self-evaluation of learning maths. However, after several RM analysis, there are only three items retained (see Table 4-4).

4.4.5.2 Math Self-Efficacy (SELFEFF)

SELFEFF was measured using nine items (e.g., '*I usually do well in mathematics'*) and used to measure confidence in mathematics. A student with relatively high levels of confidence in doing math can agree with this item. Though the variables come from Students Confident in Mathematics Scale, basically they are not confidence indicators. They are statements of self-efficacy. Return to Bandura's (1997) view on this, and confidence is a descriptor utilised to identify general feelings of self-worth.

Moreover, confidence is not specific to a domain, for example, the domain of mathematics. Self-efficacy, in contrast, is a term utilised to define a person's perceived skills within a particular domain. It is a word used to self-label individual's abilities about something as specific as mathematics or at even a specific part of that subject, such as geometry or algebra. Based on this explanation, the scale used for this study is better identified as math self-efficacy (SELFEFF) than confidence.

Table 4-4 shows the result of self-beliefs measurement using RM in *Winsteps*. Based on the goodness of fit mention before, the fit of the two-scale is in the acceptable range (Linacre, 2011a).

Table 4-4: Fit statistic of self-beliefs scales

Scale	1st contrast loading	Infit (mean)		Outfit (mean)		Reliability		N item
		MNSQ	ZSTD	MNSQ	ZSTD	item	person	
SELFCONCEPT	1.6	0.99	-0.7	0.9	-0.4	0.99	0.72	3
SELFEFF	1.4	0.96	-2	0.93	-2.7	1	0.63	3

4.4.6 Variable used in the study

4.4.6.1 Mathematics achievement

Table 4-5 shows the summary statistic of math and its cognitive domains. The value presented was calculated using all plausible values for each score by using IEA IDB Analyzer software followed by SPSS. To give an idea of Indonesian student compared to other countries, Table 4-5 shows benchmark, as a way of interpreting the scaled achievement results. Four points on the mathematics achievement scales were identified (Mullis, Martin, et al., 2012) as international benchmarks—Advanced (625), High (550), Intermediate (475), and Low (400).

Table 4-5: Summary statistics of mathematic achievement data

Math and its domains	N of cases	Sum of TOTWGT	Mean	Std. Dev
MATH	5795	3710311	385.84	83.93
KNOW	5795	3710311	377.73	94.97
APPLY	5795	3710311	384.31	86.58
REASON	5795	3710311	387.57	86.49

4.4.6.2 Self-beliefs

After both scales were constructed, Table 4-6 shows descriptive statistics of the scales. The standard deviation of the scale is considerably large, for SELFCONCEPT, even larger than the mean. It means that Indonesian student self-beliefs are more spread out, and therefore indicating the challenging of measuring these two self-beliefs.

Table 4-6: Descriptive statistics of Indonesian students' self-beliefs

Self-beliefs	N of cases	Mean	Std. Dev
SELFCONCEPT	5795	10.24	11.59
SELFEFF	5795	10.12	8.84

4.4.6.2.1 School climate

Next, Table 4-7 shows descriptive statistics of all school climate factors.

Table 4-7: Statistics descriptive for school climate factor scales

School climate factors	N	Mean	Std. Dev
Safety			
Student safety	5795	9.95	4.77
Teacher safety	5795	6.17	11.39
School safety	5795	8.82	8.57
School discipline	5795	5.31	6.37
Relationship			
Student like school	5795	13.64	7.54
Teacher connected with the school	5795	7.38	6.11
Teacher interaction	5795	9.62	5.62
Teaching-learning			
Student engagement with the math lesson	5795	9.17	6.42
Teacher emphasis on academic success	5795	9.28	6.33
Teacher instruction to engage student	5795	9.63	7.57
Teacher: classroom disturbance	5795	10.00	5.26
Teacher confident in teaching	5795	13.51	4.35
School emphasis on academic success	5795	10.57	5.58
Physical environment			
Teacher working condition	5795	9.88	3.15
School technology support	5795	9.31	7.50
School general resources	5795	9.57	3.48
School improvement process			
School leadership	5795	6.36	7.14

4.4.6.3 Other variables

Following the theoretical justification of explanatory variables that potentially would be chosen for inclusion in the fixed part of the models depending on statistical significance, the other variables, in addition to school climate factors, tested were categorised in four types: (1) student's characteristics (2) Teacher characteristics, and (3) School context.

4.4.6.3.1 Student's level characteristics

It has been argued that school effectiveness research has to take into account student background factors in investigating the variation of school performance (Aikens & Barbarin, 2008; Reynolds et al., 2014; Sammons, 1999). Some previous research (i.e., Mortimore, 1988; Sammons et al., 1993; Thomas et al., 1997) consistently demonstrated that some of the background factors such as prior attainment, social-economic status (SES), parental employment status, family income, and gender (Mortimore, 1988). Prior attainment is seen as

the most important factor to control in assessing the school effect (Mortimore, 1988; Thomas, 1998). However, this research, since using TIMSS database that is cross-sectional, for this reason, unfortunately, cannot include prior attainment. Consequently, the findings of the analysis need to be interpreted with cautious.

Nevertheless, there are some background data available for student and school level that can arguably be used as proxy measures for prior attainment. It is argued that these background variables are adequate alternates for prior student attainment when there is no prior attainment provided as on the TIMSS and PISA (Lenkeit, 2013; Thomas & Mortimore, 1996) such as, when assessing primary school students that only limited prior achievement can be obtained (Teddlie et al., 2000). Therefore, this study will only use background data as a means to approximate longitudinal student progress (value-added) measures because they are the only data available. The background data that will be used in this study as controlling variables are mentioned below.

Gender

Gender is one attribute of the student that may affect the results of learning (Bacharach et al., 2003; Brunner et al., 2008; Hergovich et al., 2004). Gender is a categorical variable in TIMSS 2011 that shows whether a student is a girl (1) or a boy (2). Table 4-8 shows boys' and girls' samples in the TIMSS 2011 study.

Table 4-8: Summary statistic of student's gender

Gender	Frequency	Percentage
Girl	2972	51.3
Boy	2823	48.7
Total	5795	100

Student SES

The importance of SES has been validated by many studies (Aikens & Barbarin, 2008; Felner et al., 1995). Student's SES in TIMSS 2011 is a composite variable that comprises three measurements, including a) parents' education, b) the number of study support, and c) the number of books. Table 4-9 shows the Indonesian student' SES.

Table 4-9 Summary statistic of student's gender

SES indicators	Frequency	Per cent
Parents' education		
University or higher	712	12.3
Post-secondary but not university	403	7.0
Upper secondary	2052	35.4
Lower secondary	1236	21.3
Some primary, lower secondary or no school	1392	24.0
Total	5795	100.0
Number of books		
0-10 books	1425	24.6
11-25 books	3080	53.1
26-100 books	1059	18.3
101-200 books	155	2.7
More than 200	76	1.3
Total	5795	100.0
Home study support		
Neither own room nor electronic goods	880	15.2
Either own room or electronic goods	3983	68.7
Both own room and electronic goods	932	16.1
Total	5795	100.0

Language

Language is a proxy measure for ethnic background. As explained in Chapter 2 that Indonesia comprises of many ethnics and languages. The language variable is about the use of language testing (*Bahasa Indonesia*) in daily student life. This categorical variable measure student's frequency in using the language (Table 4-10).

Table 4-10: Summary statistic of the frequency using Bahasa Indonesia

Language use	Frequency	Per cent
Always	2172	37.5
Sometimes	3226	55.7
Never	397	6.9
Total	5795	100.0

4.4.6.3.2 Teacher characteristics

Previous research suggested that teachers do make a difference in student achievement (e.g., Darling-Hammond, 2000b; Heck, 2009; Sanders et al., 1997). Teacher's characteristics variables that informed by teacher effectiveness research were included as explanatory variables at the teacher level. There are four teacher characteristics available in the TIMSS 2011 teacher background questionnaire, there are: teacher gender, age, the teacher's educational background, and teacher's teaching experience.

a. Gender

Table 4-11 shows the variable of teacher gender. About 57.2% of the teacher are female, and 42.8 are male.

Table 4-11: Teacher gender

Gender	Frequency	Per cent
Female	3317	57.2
Male	2478	42.8
Total	5795	100.0

b. Age

About 65% of the teachers are in age between 30 to 49 (See Table 4-12).

Table 4-12: Teacher age

Age	Frequency	Per cent
Under 25	286	4.9
25-29	813	14.0
30-39	1479	25.5
40-49	2366	40.8
50-59	826	14.3
60 or more	25	.4
Total	5795	100.0

c. Educational background

Table 4-13 shows teacher's education background and can be divided as major in math education, all other majors, and no formal education beyond upper-secondary.

Table 4-13: Teacher educational background

Teacher education	N	Percentage
Major in math	155	91
All other major	13	8
No formal education beyond upper-secondary	2	1
Total	170	100

d. Teaching experiences

Table 4-14 shows the math achievement differences based on how long a teacher has been teaching.

Table 4-14: Teaching experiences

Teaching experiences	N	Percentage
20 years or more	49	29
10 - 20 years	56	33
5 – 10 years	35	21
Less than five years	30	18
Total	170	100

4.4.6.3.3 School context

School contextual factors relate to the general school environment and thus are beyond the control of the school staff. The context is a given condition that has to be considered (De Fraine et al., 2002; Opdenakker & Van Damme, 2005; Sammons et al., 1994) as students are not randomly allocated to schools and are influenced by socioeconomic and ethnic residential segregation and other non-random practices. According to Teddlie and Reynolds (2000), there are four context characteristics has taken into account: SES of the student body; location (urban/rural); grade phase of schooling and school governance structure. Following these suggestions and the availability of the data in TIMSS, school's SES, school location (urban-rural), school types (general-Islamic) were chosen as school context variables.

Additionally, school location and school types also have a noticeable impact. School location refers to where the school is located, rural or urban (Young, 1998), while school types refer to differences in curriculum application. Indonesia has two types of schools, namely the Islamic school and general school. These types of school are considered an important variable because the differences in their practices might affect school climate (Sikkink, 2012).

a. Average SES

School average SES is an aggregated score of student's SESs. Table 4-15 shows a description of the average SES.

Table 4-15: Summary statistic of average student's SES at the school level

Student SES	N	Mean	Std. Dev
SES	153	8.53	0.78

b. School type

As explained before, the sample schools were explicitly stratified by public and private schools as well as general and Islamic schools. The sample was chosen based on the dual schooling system in the Indonesian context (Ministry of Education and Culture, 2012). Table 4-16 shows the school type in the Indonesian context.

Table 4-16: School types

School types	Frequency	Percentage
Public General	88	58%
Public Islamic	8	5%
Private General	32	21%
Private Islamic	25	16%
Total	153	100%

Source: Methods and Procedures in TIMSS & PIRLS, 2011 (Joncas & Foy, 2012).

c. School location (urban-rural)

To present data about the school's location, TIMSS 2011 asked headteachers to define the population size of the city, town, or area in which their schools were located. Table 4-17 presents the detail of the location of the sampled school.

Table 4-17: School location

School's locations	Frequency	Percentage
Urban	10	7%
Suburban	85	56%
Medium size city	15	10%
Small town	36	24%
Remote rural	7	5%
Total	153	100%

4.4.7 TIMSS 2011 data analysis plan: Multilevel model (MLM)

The purpose of the quantitative data analysis is to address RQ1 (What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?). The TIMSS database comprises of student records that are grouped within a school, and thus it provides a data hierarchy including both student-level and school-level variables. For this reason, it is possible to perform a multilevel modelling analysis (Goldstein, 2011). Besides, one of the advantages of MLM approach to data analysis is its ability to handle both unstructured and unbalanced data, as in this TIMSS data, of 153 schools, only 20 schools have two classes (Hox, 2002). MLwiN software (Rasbash et al., 2012) was used in this study. This multilevel statistical package is one of the most suitable software to conduct multilevel modelling analysis (Galecki & West, 2013).

To address research questions, MLM were used. The MLM approach has been used since the late 1980s in school matters research in the UK (Mortimore et al., 1989), then many studies followed. MLM has some benefits in order to address RQ1 of the study including: the ability to simultaneously estimate the effect of the different variables at both student, classroom

and school level (Raudenbush & Bryk, 2002), and place explanatory variables at their proper hierarchical location (Hox, 2002), to avoid aggregation or disaggregation of the data. The technique also offers correct standard errors, confidence intervals and significance tests (Raudenbush & Bryk, 2002).

The data from TIMSS is examined using a three-level hierarchical model with school at level 3, teacher/classroom at level 2, and students at level 1. Most past studies have employed a 2-level model (school at level 2, and students at level 1). However, the author considered the difference between the general and Islamic types of school in Indonesia should not be taken for granted. As described in Chapters 1 and 2, students in these two types of schools experienced different classroom circumstances. This makes it imperative to learn more about how the classroom experience affects school effectiveness. Moreover, some researchers also stressed that the classroom level is a critical determinant of school effectiveness, especially by considering that learning mainly takes place in a classroom (Chapman et al., 2012; Hill & Rowe, 1996; Rowe & Hill, 1998). Instead of using the common approach of using 2 hierarchical models (i.e., school and student levels), the present research considered classroom as an additional important level.

Moreover, theoretically, this data set can be evaluated using a 3-level model, because the multi-level model can accommodate unbalanced data (Hasselmann, 2018; Jones, 2019) as in TIMSS 2011 Indonesia dataset, where there are only 20 schools with more than one classroom participating in the study.

Different statistical models was conducted. The first was the null model, estimated the total variance and its components. This analysis aimed to estimate overall mean achievement. Besides, to see whether there were any school differences and classroom differences in mean achievement. A stepwise procedure was conducted, where a group of related variables were entered at a time. It starts from the simplest models building up to more complex models. Any

predictor that did not have a significant contribution to the learning outcomes variance was omitted to refine the final model.

Model 1 was extended by adding explanatory variables measured at the student level, then classroom level, and school level. The purpose of fitting that model was to find out which of the pupil, class and school factors had a significant effect on student achievement. Then followed by another model as described in table 4-18. Then, statistically significant variables were added together as a group to see their relevance in explaining the students' learning outcome. Overall, the analysis was conducted into several steps, in each case, both maths achievement and self-beliefs measures were employed separately as outcome measures (see Yu and Thomas (2008) for comparable steps).

Table 4-18: Research questions and multilevel models

Research Questions	Model employed
RQ1.1: What is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?	Model 0/Null Model: No explanatory variables
RQ1.2: After controlling for student, teacher, and school characteristics, what is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?	Model 1: Explanatory variables comprises student background characteristics (SES and Gender) Model 2: Explanatory variables comprise teacher characteristics (education background and teaching experiences). Model 3: Explanatory variables comprise school context variables (Average SES, location, and school types). Model 4: Combine Model 1,2, and 3
RQ1.3: What are the school climate factors that significantly explain the variance between school and classroom performance among Indonesian Year 8 students in math and self-beliefs before and after adjusting the characteristics of the student, teacher, and school?	Model 5: school climate factors only (individually and aggregated at school or classroom level) Model 6: combine 1,2,3,5 (final model)
RQ1.4: Do students in general school have higher achievement and self-beliefs than *madrasah* before and after controlling school climate and other factors?	Compare between Model 4 and Model 6

4.4.8 Limitations of MLM

Technical concerns to address under MLM are accuracy and the retrospective nature of data as well as measurement errors. These issues need to be recognised in awareness when interpreting the findings (Sammons et al., 1997). In cases like the present study, in which the data is from cross-sectional data, a lack of prior attainment measures and only one to two classes available at one school, the generalisation of its findings should be cautious and tentative.

Imprecision, despite the type of assessment used for measuring student achievement, ambiguity will always be a problem (Myers & Goldstein, 1998; Rosenkvist, 2010). Given that imprecision can be reduced through publishing student test results with a margin of error, this margin tends to be so large in MLM that this approach is useful for identifying schools at the extreme of a performance distribution, but not with the precision needed for rank ordering schools.

4.5 Qualitative phase

Most of the statistical analyses, in particular MLM are compatible methods to detect variances among schools. However, the results do not reveal the reasons or explanations regarding the variances found in school performance that might arise. Therefore, the quantitative phase was followed qualitative phase by using illustrative cases. The case study was chosen because this type of qualitative method helps to intensify, describe and analysis more holistic of a bounded phenomenon in an institution (Merriam, 1988) like school, as the case in this research. It is also can be used to verify and explore the existence or absence of certain circumstances, processes or attributes related to school climate for different schools in terms of their academic performance and curriculum implementation. In this way, the researcher used the cases (of different schools) to explore in-depth school climate differences and then compare them.

4.5.1 Using illustrative case study schools

The illustrative case study school is selected to understand something else, school climate differences and or similarities among the schools in this study (Stake, 2006). By using a case study, the researchers aim to get understanding of people's perception about their experiences (Stake, 2006; Yin, 2014). In the case study research, the selected cases were believed to have fruitful findings relating to the research question (Grandy, 2010). In this research, four case study schools with different characteristics were selected. The selection aimed to get a fruitful explanation of how and why specific factors of school climate examined in the first phase were significant or insignificant to explain students' self-beliefs and achievement.

By connecting estimations of school performance resulting from MLM analysis, the application of how school climate is practised in particular schools can be observed and examined by the case study. So, the case study focused on the school practice of selected schools in term of their school climate. This "field-based" data will form far more deep evidence (Yin, 2014), and compare them with the finding with the "desk-based" data in the first phase.

4.5.2 Data collection techniques and procedures

4.5.2.1 Piloting

In the qualitative phase of the study, the data was collected between September and November 2015. Before employing the main study and using the qualitative interview instruments in the actual sample, the tools consisting of the interview and FGI questions were piloted.

Piloting is a procedure of preparation before conducting the main study (Tashakkori & Teddlie, 2003). It aims to address potential functional and practical issues in the research procedures, whether the structure is appropriate for the main study (Seidman, 2006; Silverman,

2013) as well as trying out the questions. A pilot study also useful to identify if there are any errors or restrictions within the interview design. Then, it allows modifications or adjustment to the interview questions before embarking into the main study (Silverman, 2013).

Piloting was conducted at *SMP Negeri 3* Lubuk Pakam (lower secondary school - general public school) in Deli Serdang District, North Sumatera. This piloting study aimed to make sure that the interview protocol can be delivered and to give an idea about the flow and other technical matters during the data collections. One FGI of the teacher (8 teachers), one FGI of the student (8 students), and one interview with headteacher were conducted.

There were two main revisions to the instruments based on piloting results: (1) there was a need to adjust some questions to be more understandable, for example, the need of explaining of the school climate term. In the pilot study, the school climate was interpreted as the physical climate of the weather like hot or cold weather. (2) The interview protocol could not run strictly because of time restriction and the flow of questions. The maximum time allowed was only 40 minutes. Regarding the flow, when asking a question on one school climate factors, the participants sometimes related it with other factors. Therefore, if the researcher believed that other factors have been covered in the previous question, then he jumped to other questions that have not been asked.

4.5.2.2 Data collection techniques

This research used interview and Focus Group Interview (FGI) to collect data in the second phase (qualitative) study. Interview and FGI are common data collection methods in qualitative research (Gill et al., 2008). The semi-structured interview contains key questions - as in the theoretical framework used – and allows the researcher to probe flexibly on particular issues by asking follow-up questions. Unstructured interviews do not reflect any predetermined theories or ideas and are performed with the little arrangement (Gill et al., 2008). It aimed to

explore the policymaker view about a certain research topic. It started with a general opening question for instance ‘Can you tell me about policy in creating a positive school climate?’

On the other hand, FGI were used for gathering data on collective views and aimed at generating a perception of participants’ experiences in the school context (Morgan, 2007). There are three techniques used for collecting the data. (1) In-depth semi-structured interviews with headteachers; (2) focus group interview (FGI) with a group of teachers and groups of students in each school. (3) Freestyle/unstructured interview with policymakers (one in MORA, 2 in MOEC, and a local policymaker).

The written transcripts were used to recode the information collected during the interviews and FGI. The participants were asked for written consent to record their interview responses electronically. The interview protocols were piloted, and the final version will include ten-fifteen open-ended questions based on the results of phase 1. The questions were designed to address Thapa et al. (2013) five school climate factors as well as the effectiveness evaluation dimensions adapted from the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) and make sure that the necessary dimension is covered to explore how school climate effects on students' academic outcome and self-beliefs. The questions were developed based on the combination of the two theoretical frameworks adopted in this study. The complete questions can be seen in Appendix 8.

4.5.2.3 Participants

4.5.2.3.1 Student, teacher and headteacher

The participants of this qualitative phase of the study included students, teachers, and headteacher. Teachers and students were grouped for interviews to make focus group interviews, whereas the headteacher was interviewed face-to-face. Table 4-20 describes the participants in each school.

Table 4-19: Participants in each school

School	Location	Headteacher interview	Teacher's FGI	Student's FGI
Madrasah A	Medan City – North Sumatera (urban)	Former vice Headteacher (Male)	2 teachers (All Female)	7 students (4 Female, 3 Male)
Madrasah B	Semarang City – Central Java (urban)	Headteacher (Female)	3 teachers (1 Female, 2 Male)	8 students (All Female)
School C	Tanjung Morawa City, Deli Serdang – North Sumatera (rural)	Headteacher (Female)	3 teachers including the headteacher (2 Female, 1 Male)	8 students (4 Female, 4 Male)
School D	Padang City – West Sumatera (urban)	Headteacher (Male)	6 teachers (3 Female, 3 Male)	8 students (4 Female, 4 Male)

4.5.2.3.2 Policymakers

In addition to qualitative data collected from illustrative case study schools, interviews with four policymakers were also conducted to obtain further evidence on school climate policy. The idea was to get some information about how the local and national government policy influence schools in creating school climate. The participants included:

1. Head of Education Office in Padang City, West Sumatera
2. Ministry of Religious Affair (MORA) in Jakarta: Head of Curriculum and evaluation, Directorate of *Madrasah* Education
3. Ministry of Education and Culture (MOEC) in Jakarta:

   a. Head of curriculum and evaluation, Directorate of Junior high school

   b. Head of Facilities affairs, Directorate of lower secondary school

4.5.2.4 School selection procedure

Four schools were selected to participate in the case study. The selection based on schools' performance in mathematics in TIMSS 2011 study, as well as their school climate. To classifying the school performance, schools were grouped as described in Table 4-21.

Schools were ranked based on those classifications. Then, schools on percentile 25 classified as poor performance schools, and on percentile 75 classified as good performance

schools. Next, because this study also aims to compare general and Islamic schools, one good and one poor performance schools were selected based on their school types. Therefore, in total, four schools would be included as illustrative cases. Other considerations were the location and accessibility of school. Then the researcher sent letters to ask for their participation in the study.

Table 4-20: School classification

Percentile	Maths score	School Climate	Math Self-Concept	Classification
25	< 349.07	< 9.21	< 8.03	Poor performance school
50	349.07– 425.25	9.21 – 10.82	8.04 – 11.90	Average School
75	> 425.25	10.82	> 11.90	High performance school

The final schools that participated in the study are described in Table 4-22. The school profiles, as described in Table 4-22, reflected the analysis of TIMSS 2011 in the first phase of the study. These results were also consistent with BAN-SM's (National Accreditation Board of School/*Madrasah)* report of schools' accreditation. Low-performance schools as analysed in the first phase of the study were also judged by BAN-SM as need improvement schools, indicating that the schools are in the lowest standard of schools in Indonesia and vice versa.

BAN-SM is an independent institution that assesses the feasibility of primary and secondary school regarding formal national education standards (Ministry of Education and Culture, 2005). Data from this board is considered essential because the accreditation process also measures the effectiveness of the school in promoting students' learning. They measure eight aspects of the school including (1) curriculum content; (2) teaching-learning process; (3) pupil competencies; (4) teacher and school-staff competences; (5) school facilities; (6) school management; (7) school funding; and (8) educational assessment. As based on the Indonesia Education Act Year 2003, schools need to be accredited every five years. For that reason, by using the accreditation result, the selection of school can be more rigorous because it can acquire a recent portrait of the school. So, the school selection is based on school performance measures from TIMSS and estimates of overall school effectiveness by BAN –SM.

Table 4-21: Profile of participated schools

Schools	Maths Score (Raw)	ASC	School Climate	Rank residual*	Accreditation Report**
Madrasah A	314.44	11.30	10.60	29	C
Madrasah B	460.67	9.98	11.44	119	A
School C	343.37	10.89	10.32	23	C
School D	497.54	9.92	12.38	148	A

*After controlling significant school background variable on math score of 153 schools: location, type of school (Islamic and general), public and private.
** A = Outstanding, B = Good, C = Need Improvement. Information obtained from *Badan Akreditasi Nasional Sekolah*/*Madrasah* BAN-SM (National Board of Accreditation for School and *Madrasah*)

4.5.3 Qualitative data analysis

There are plenty of ways and variation method to analyse qualitative data. At least 50 different types of analysis to choose (Wolcott, 1994). For instance, content analysis approach (Anfara Jr et al., 2002; Hsieh & Shannon, 2005), interpretative phenomenological approach (Moustakas, 1994; Smith et al., 2009), template analysis approach (King, 1998), discourse analysis (Brown et al., 1983), narrative analysis (Riessman, 1993), conversation analysis (Ten Have, 2007), and thematic analysis (Braun & Clarke, 2006). To all the methods of analysis mentioned, Swain (2018) concluded that the key principle of qualitative data analysis is reducing data, displaying, and drawing a conclusion as summarised by Miles et al. (2014).

Of the method of analysis, in the social sciences, a thematic analysis approach is the most popular qualitative analytic approach (Swain, 2018). Braun and Clarke (2006) argued that thematic analysis is a flexible qualitative analysis approach that can be employed regardless of the researcher's philosophical position. This approach is a fundamental analysis method that researcher must learn since the thematic approach is providing foundational key skills and techniques that are used in other approaches of qualitative analysis (Braun & Clarke, 2006). This method of analysis mainly aims for identifying and coding patterns (Braun & Clarke, 2006) and searching themes (Fereday & Muir-Cochrane, 2006) that link with a specific research question.

Thematic analysis was chosen in this qualitative phase of the study, but in a flexible form, a hybrid approach of thematic analysis (Fereday & Muir-Cochrane, 2006; Swain, 2018). First, this approach exercised a-priori codes which were developed before examining the current data derived from existing analytic framework selected in the study (see literature review). Second, followed by an inductive approach or data-driven allowing for themes to emerge from the collected data.

This process of analysis was carried step-by-step. First, the data collected was transcribed into a word processing document, and then all the transcriptions were deposited together into NVivo. The researcher classified the data into meaningful analytical units using the theory-driven a-priori codes (Thapa et al., 2013) and its effectiveness evaluation dimensions (Creemers & Kyriakides, 2008) (see the literature review). This procedure was made through reading, listening and summarising the raw data. Then the data-driven or inductive codes were additionally developed.

4.5.4 *A-priori* themes: 5 school climate factors combined with effectiveness evaluation dimensions

A priori themes allow researchers to capture key factors that have informed the design and aims of a study and to adopt assessment criteria that a research project has been planned to address (Brooks et al., 2015; King, 1998). This research uses a-priory themes as research analytic framework to inform the analyses.

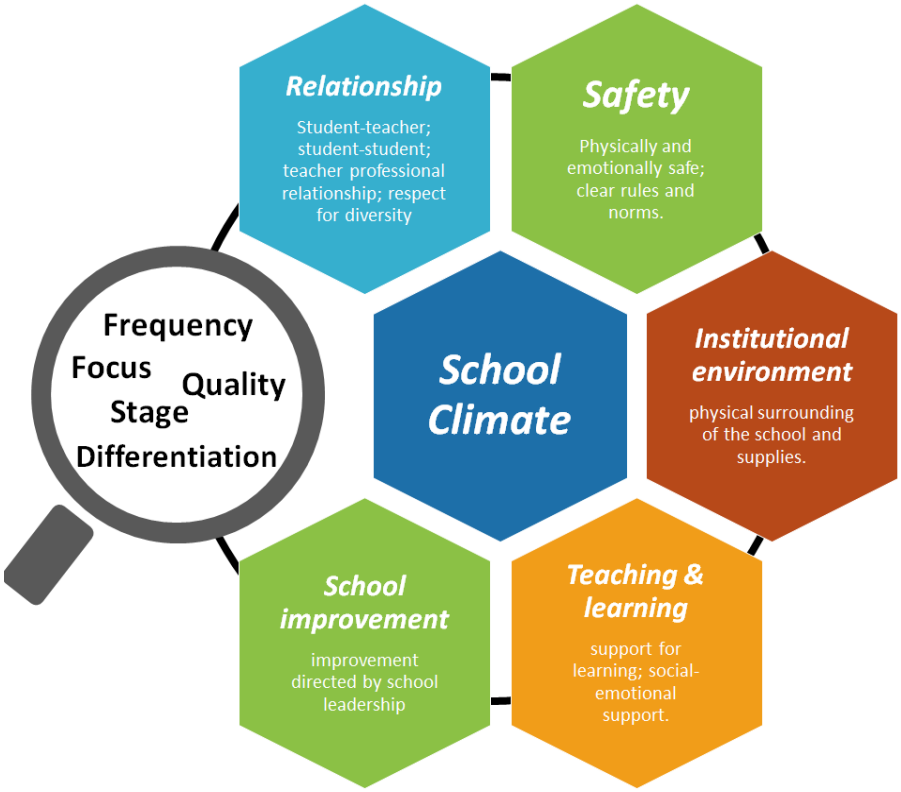


Figure 4-3: School climate effectiveness model

As explained in Chapter 3, there are five factors of school climate (Thapa et al., 2013) and five DMEE's evaluative dimensions (Creemers &Kyriakides, 2008) that will be used to as an analytic framework to guide the identification of codes key and themes. So, each school climate factor, as reviewed by Thapa et al. (2013) have their own effectiveness evaluation dimensions. First, frequency refers to the quantity that an activity associated with a school climate factor happens or is present in the school or classroom. The other four evaluation dimensions (focus, quality, stage, and differentiation) examine qualitative characteristics of the functioning of the school climate factors and help to describe the school climate. However, it should be noted that not all the evaluation dimensions can be applied if, for example, participant responses provide insufficient relevant evidence in interviews or focus groups.

By combining these concepts, it aimed to explore and evaluate the overall quality of school climate factors in a systematic way. So, the whole picture of school climate might be captured, and this approach may give more information to the policymakers, headteachers,

teachers, and students how to develop and evaluate school climate. Figure 4.3 represents the combined model of the school climate factor and its effectiveness evaluation dimensions.

This analytic framework is exercised as the preliminary codes in order to look for ways in which the data fitted or did not fit these codes. First, the researcher thoroughly read, and the transcriptions of the interviews compared the information obtained in each school in reference to each of the school climate effectiveness model. Reporting disconfirming evidence (for instance, not all effectiveness evaluative dimension can be captured) was also employed to improve the accuracy of the data analysis. Creswell and Plano Clark (2011) suggested that expectation of evidence for themes in real life are more than just positive information. Secondly, to make sense of the data that was not included in the *a priori* codes, the researcher took memos about confusing, insightful, and interesting aspects of the data. At the same time, the researcher was taking note of potential new themes that may emerge from the data. As follows, these new themes/factors were necessarily needed to better describe the local context of the study. Finally, representative quotes were selected that best-represented participants' perspectives on the research themes.

4.5.5 Data-driven categories

It was found that the *a priori* codes were relevant when describing the school climate practices but should be adapted to the local context. Concerning the key role of the context when reviewing school climate, the analysis challenged the need to expand this framework by subsequently including other factors that helped to explain new patterns that emerged from the data. In searching for new themes, the interaction of text, codes, and themes in this study involved some repetitions before the analysis proceeded to a useful and informative phase. Then, themes were further grouped and assigned concise words to label the sense that highlighted the theme (Fereday & Muir-Cochrane, 2006).

4.5.6 Strengths and limitations of the qualitative approach adopted

Since the analytical process of qualitative data employed a well-known theoretical framework, starting the analysis with *a-priory* factors/themes in the case study was a step forward when attempting to build on prior research. It recognises that further research is required that replicating, clarifying and building on the results of previous research. Therefore, the purpose of this qualitative analysis was not to create an additional list of school climate factors, but to employ a well-known concept to a new context. This decision was also made because the school climate factors to be a meaningful category that assisted interpretation, analysis and identification of issues contained in the data. It also facilitated the researcher to communicate the findings and movement toward a more analytical level.

This analytical approach provided useful guidelines when describing the similarities and differences between schools. Complementing the analysis with a data-driven inductive approach provided a relevant tool for making sense and exploring the tangled complexity of the context in which the schools were operating that was not covered with the a-priori codes. However, choosing this approach was not made without its problems.

The main research instrument was an interview schedule intended to elicit the participants' opinions on the combined framework of school climate and effectiveness evaluation dimensions, so it was not capable of gathering systematic data on other aspects. There was also an imbalance with more focus on the deductive coding (derived from the theoretical framework) than the inductive coding (themes emerging from the raw data). Finally, the new factors that emerged from this study incorporating additional school climate factors could be potentially used in the future as an evaluation checklist to identify areas of strength and weakness, but not as a final blueprint or guarantee for school climate, which is beyond the scope of this study. This emerging framework can illuminate future research carried out in complex social contexts such as Indonesia and maybe other similar contexts.

4.5.7 Trustworthiness

The criteria to judge a good qualitative study is dissimilar from the quantitative one. Trustworthiness has become a key idea in qualitative research because it facilitates researchers to identify values beyond criteria commonly used in quantitative research (Given & Saumure, 2008). Trustworthiness can be evaluated through a verification process rather than a traditional test of reliability or validity. The uniqueness of the qualitative analysis in a particular context prohibits it from being exactly replicated in a different context. However, statements concerning the position of the researcher such as the central assumption, informant choice, perceptions and beliefs may increase the possibility of reproduction of the study in a different context (Creswell, 2003).

To confirm the findings or, in other terms, to determine whether the evidence is accurate and whether it refers to reality (Merriam, 2002), three key procedures will be used in the qualitative phase. (1) triangulation – linking various information sources (interviews and FGI); (2) providing a rich comprehensive summary of the findings; and (4) external audit – allow an outside person to perform a detailed research review and give feedback (Creswell, 2003).

4.6 Ethical issues

Ethics refers to a moral system, where others can judge an action as right or wrong. For the researcher, it should be a case of professional integrity to enhance ethical research (Denscombe, 2010). Since this research will deal with the amount of personal data, ethical concern (BERA, 2011) was address as follow:

4.6.1 Data protection

This research used two types of data, quantitative and qualitative. Quantitative data was obtained from TIMSS 2011 dataset. Although the TIMSS data is open to the public, it is still necessary to keep the data securely, because the analysis findings may reflect some schools' performance. Therefore, the data was saved in the researcher password-protected computer.

Next, since the data from the first phase informed the second phase of the study, it is important to get written permission for accessing school IDs from MOEC or local government authorities, and this has been granted.

Then, qualitative data was used in the second phase of the research. The researcher is aware that the qualitative data include sensitive information regarding schools' and students' private data. For this reason, the data was stored carefully to make sure that neither school nor individual can be identified in the findings. Students' name, teachers', as well as the headteachers in this research are all anonymous by referring to them using codes. Participants were notified of this policy through an inform consent form. Also, the data then was stored in the researcher's password-protected computer before storing it in the University of Bristol 'O' drive. In addition, anonymity and confidentiality of participants and the school have been guaranteed through the whole research process.

4.6.2 Informed consent

For qualitative data, written consent for interviews and permission to quote the interviewee's response have been acquired from all schools and participants (students, teachers, and headteachers). The process will be done by discussing and asking them to sign an informant's consent form. Since all students that were interviewed are under 18, both the consent of their own and their parents were obtained. Before the data gathering starts, the researcher also made clear to all schools and participants that they have the right to withdraw at any time.

4.6.3 Researcher's position

It is obviously important to take into account and aim to reduce researchers' bias because no research approach is impartial and has an impact on the way that the data are analysed (Hammersley & Atkinson, 2007). Researcher's identity as a Moslem, when researching a specific school (in this research, Islamic schools - *madrasah*), might cause bias

in some cases. Though, it has been assisted in explaining the school in more depth. For that reason, the researcher has considered the school from a distance and keep the position as an outsider.

4.6.4 Other moral concerns

The identification as a good or bad school climate might risk negative effect on selected schools. Also, access to teachers and students, although it has been negotiated with head-teachers, might interfere with school activities. Hence, the researcher has implemented a positive and respectful attitude to every school. In particular, the researcher has not informed or labelled a school as good or bad but instead saw them as room for development to provide better learning for all students. This moral concern has not only ensured the ethics of the study but also facilitated the process of data collection by creating trust and mutual respect between the researched and the researcher (Hinkin et al., 2007).

4.7 Chapter summary

This chapter began by recognising the ontological and epistemological assumptions of the study because any field of investigation is inherently related to a particular paradigm. Pragmatism was chosen as the philosophical bases of this research. From an ontological perspective, the nature concept of school climate is known not only to be objective but also to provide a comprehensive understanding of the school climate. From the epistemological point of view, the researcher believes that in acquiring knowledge, it is more important to highlight methodological approach that fits well with the specific issue of analysis, since knowledge construction is contextual in nature and influenced by the cultural, political, and historical conditions.

The data collection and analysis appeared in two consecutive phases: the analysis of secondary data in the quantitative phase and the collection and analysis of the follow up qualitative data. In the quantitative phase, MLM was used to conduct a secondary data analysis

of TIMSS 2011. The MLM employed a model building procedure containing several models sequentially. Then, to conduct qualitative phase four illustrative school cases were selected. To emphasise the qualitative analysis, a hybrid approach of thematic analysis was adopted. It began by applying as deductive *a priori* codes of school climate factors and then followed by a data-driven inductive approach letting for themes to arise directly from the data. Lastly, the ethical issues were emphasised and need to be counted when interpreting the findings and conclusions of the research presented in the following chapters.

Chapter 5: Quantitative Findings

5.1 Introduction to the chapter

This chapter presents the findings of the first phase of the research. Specifically, the author reported the quantitative analyses of the Indonesian TIMSS 2011 data. The chapter aims to answer the third objective, RQ1 and its sub-questions by applying MLM techniques. Specifically, the author analysed the differences between schools in terms of mathematics performance and self-beliefs among Year 8 Students in Indonesia. The author presented the results based on the sequence of the proposed research questions. Subsequent to answering the five sub-questions of RQ1, this chapter concluded by highlighting the main findings of the first phase of the research.

5.2 RQ1: Difference in mathematics and self-beliefs

RQ1.1 asked, "What is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?" To answer this question, the researcher investigated the variances in math performance and self-concept based on school-level and classroom-level. Specifically, by employing the variance components model for a three-levels model (school, classroom, student levels). Model 0 (see Table 5-1) describe school performance of the Indonesian secondary schools in terms of mathematics score (MATH), mathematical Knowledge (KNOW), ability to apply mathematics (APPLY), and mathematical reasoning (REASON). The table also describe the affective components, which includes student's self-concept in mathematics (SELFCONCEPT) and their self-efficacy in mathematics (SELFEFF), which is indicated by the proportion of variance explained by each level.

The results showed that Indonesian students' mathematical performance varied among schools. Specifically, the school-level explained about 35% of the variance of MATH, 36% of

the variance of KNOW, 34% of the variance of APPLY, and 26% of the variance of REASON. The predictive power of the school-level was the lowest in terms of REASON. This may be attributed to the exam-oriented learning practices in Indonesia, which puts little emphasis on the capacity of logical thinking and the ability to conduct intuitive and inductive reasoning to reach a solution that were measured by TIMSS (Berkhout et al., 2019; Furaidah et al., 2015; Ragatz et al., 2015).

The school-level explained a magnitude of variance on the cognitive domains of mathematics. The classroom-level had lesser predictive power. Specifically, the classroom-level accounted for only about 9% of the variance of MATH, 8% of the variance of KNOW, 8% of the variance of APPLY, and 5% of the variance of REASON. This finding indicated that Indonesian students' mathematical performance varied less among classrooms. Caution is warranted to apply these results because of the small sample of classes in the TIMSS dataset. As mentioned in Chapter 4, from the 74 schools that are available in the TIMSS dataset, only 20 schools with more than one classroom. As such, there might have been not enough power to conclude the differences in math performance at the classroom-level.

In terms of the affective domain, the school-level accounted for only about 5% of the variance of SELFCONCEPT and 10% of the variance of SELFEFF. Similarly, the classroom-level accounted for only about 4% of the variance of SELFCONCEPT and 6% of the variance of SELFEFF. These results were consistent with past studies that have shown a small variation of non-cognitive outcomes when comparing one school to another (Gray, 2004; Thomas et al., 2010).

The variation of mathematics performance was best accounted for at the student-level. The student-level predicted more than 56% of the variance of MATH, 58% of the variance of KNOW, 58% of the variance of APPLY, and almost 70% of the variance of REASON. The student-level also accounted for 91% of the variance of SELFCONCEPT and 84% of the

variance of SELFEFF. As expected, these results were consistent with past studies that showed individual student-level played a major role in predicting students' performance (Gray et al., 2001; Thiele et al., 2016).

A comparison of the empty models of all outcome measures revealed that the differences between schools were more noticeable on academic achievement than in self-beliefs. This findings are in line with the results of studies conducted in Belgium (Opdenakker & Van Damme, 2000), the UK (Gray, 2004; Thomas, 2001), and Cyprus (Creemers & Kyriakides, 2010a). These studies also found that differences between schools in terms of affective outcomes were smaller in comparison to the results of the academic outcomes.

5.3 RQ1: Difference in mathematics and self-belief controlling for school stakeholders' characteristics

RQ1.2 asked, "After controlling for student, teacher, and school characteristics, what is the range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs?". Model 1, 2, 3, and 4 were utilised to answer RQ1.2. Each model was aimed to verify the relationship between student characteristics (Model 1), teachers (Model 2), and school (Model 3), and learning outcomes. Then, Model 4 combined Model 1 to 3 that includes all statistically significant explanatory variables at students, teachers, and school levels for all learning outcomes (MATH, KNOW, APPLY, REASON, SELFCONCEPT, and SELFEFF). The following section describes the differences between school and classroom after controlling student, teacher, and school background characteristics.

Model l (Table 5-2) was employed to verify whether there were differences in mathematics and its domains and self-beliefs between schools and classrooms within the school

Table 5-1: Null Model for mathematics, math self-concept, and math self-efficacy

Response	**MATH**		**KNOW**		**APPLY**		**REASON**		**SELFCONCEPT**		**SELFEFF**	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
cons	393.73	4.52	386.40	5.13	392.38	4.79	393.95	4.21	10.46	0.26	10.45	0.30
Random Part												
school variance	2416.22	420.32	3186.25	559.67	2501.87	440.94	1904.78	332.87	4.47	1.69	7.80	1.97
class variance	599.66	240.22	727.73	285.63	569.48	242.50	388.86	170.62	3.55	1.50	4.81	1.69
Student variance	3893.84	99.36	5002.04	183.79	4196.03	112.56	5024.20	142.11	82.20	1.99	65.92	1.60
VPCschool	0.35		0.36		0.34		0.26		0.05		0.10	
VPCclassroom	0.09		0.08		0.08		0.05		0.04		0.06	
VPCstudent	0.56		0.56		0.58		0.69		0.91		0.84	
Deviance	64924.58		66351.12		65323.03		66289.11		42240.44		41052.82	

Note: VPC = Variance Component Partition

Table 5-2: Model 1 with student's backgrounds variables

Response	MATH		KNOW		APPLY		REASON		SELFCONCEPT		SELFEFF	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
cons	417.45	8.16	408.69	8.10	414.00	8.65	409.24	7.75	9.90	0.66	8.57	0.62
Gender												
Boy	-4.69	2.92	-8.11	3.26	-3.00	2.73	-1.31	2.56	0.34	0.28	0.68	0.28
Parents' education												
Post-secondary but not university	-10.23	6.16	-9.39	8.17	-9.02	5.58	-3.04	6.26	0.30	0.62	0.78	0.51
Upper secondary	-9.47	4.04	-9.01	4.54	-8.29	4.28	-11.98	5.35	0.55	0.42	0.58	0.39
Lower secondary	-13.59	3.59	-14.12	5.16	-12.68	4.11	-10.53	5.59	0.93	0.48	0.66	0.44
Some primary, lower secondary or no school	-12.98	4.98	-13.62	4.31	-17.90	4.37	-3.82	4.01	0.30	0.48	0.65	0.46
Home study support												
Either own room or electronic goods	-8.32	3.92	-5.64	4.11	-5.90	3.53	-6.31	3.44	0.11	0.32	0.81	0.34

Both own room and electronic goods	-7.81	4.80	-4.04	5.19	-5.50	4.33	-4.58	5.28	0.53	0.56	1.52	0.41
Number of books												
11-25 books	-6.04	2.67	-5.83	2.88	-6.03	3.39	-2.03	4.30	0.49	0.32	0.94	0.29
26-100 books	2.89	2.69	5.98	4.86	-2.00	4.52	8.60	3.89	0.62	0.40	0.86	0.40
101-200 books	-5.30	6.93	-11.29	7.99	-9.05	8.77	8.77	9.23	1.98	0.97	2.72	0.79
More than 200	-3.06	11.57	9.85	11.84	-25.35	12.76	3.01	12.69	3.63	0.91	3.90	1.07
Language												
Sometimes	-1.49	3.55	-1.27	4.19	-0.46	3.12	-1.94	3.63	-0.99	0.30	-0.75	0.25
Never	-6.39	4.81	-6.01	5.72	1.35	5.43	-22.05	6.90	-2.59	0.50	-2.36	0.52
Self-Belief												
SELFCONCEPT	1.39	0.17	1.67	0.19	1.27	0.19	0.99	0.21				
SELFEFF	0.28	0.15	0.18	0.21	0.30	0.23	0.72	0.20				
Achievement	NA		NA		NA		NA		NA		NA	
MATH1									0.03	0.00	0.01	0.00
Random Part												
school variance	2310.22	402.13	2964.24	540.26	2356.94	427.47	1838.26	325.62	4.44	1.68	8.47	1.90
class variance	577.47	236.26	739.65	288.11	579.15	245.96	390.45	175.20	4.05	1.45	4.72	1.56
Student variance	3675.27	98.74	4684.36	177.38	4001.48	121.95	4776.05	154.40	77.92	1.80	64.17	1.56
VPCschool	0.35		0.35		0.34		0.26		0.05		0.11	
VPCclassroom	0.09		0.09		0.08		0.06		0.05		0.06	
VPCstudent	0.56		0.56		0.58		0.68		0.90		0.83	
Deviance	64561.60		65973.57		65048.87		65999.03		41945.31		40906.99	
School variance explained	4%		7%		6%		3%		1%		-9%	
Class variance explained	4%		-2%		-2%		0%		-14%		2%	
Student variance explained	6%		6%		5%		5%		5%		3%	
Total Variance explained	5%		6%		5%		4%		4%		1%	

Note: Gender is a dichotomous variable with the girl as a reference group; SE stands for standard error.

after adjusting for student characteristics/background. The selection of student background characteristics is based on DMEE (see Chapter 4 for detail) and the availability of the variables in the TIMSS 2011 data.

5.3.1 Student characteristics

There are three student's background variables associated with the socio-cultural aspect of a student which include gender, language (as a proxy for ethnicity), and social and economic status (SES). The student SES in TIMSS 2011 comprises: a) parents' education, b) the number of study support, and c) the number of books. Moreover, the student's psychological backgrounds (self-beliefs) were also included in the model. Student's self-beliefs were included as an explanatory variable as argued in Chapter 3, that the self-beliefs also have a significant effect on achievement and vice versa. Therefore, when the model examines self-concept outcomes, math achievement was also included as the explanatory variable for that non-academic outcomes.

Taking account for student background variables (see Table 5-2), 34 to 35% of the variance in MATH, KNOW, and APPLY is attributable to the difference between schools, and 26% in REASON. This proportion was as similar as in Model 0 except for APPLY, where school variance is 1 % lower after including the student's background variables. Moreover, differences between classes within the school were respectively similar. For self-beliefs outcomes, this proportion of school differences in SELFCONCEPT and SELFEFF also quite similar to Model 0.

In terms of the student's background characteristics, this study found that gender only has a statistically significant relationship with KNOW and SELFEFF, but in a different direction. For KNOW relationship with gender is negative, where boys have typically lower KNOW score compared to girls. Conversely, the boys have a higher SELFEFF score than girls.

On other student's socio-cultural factors, parent education and the number of books were also found to have a significant relationship with student achievement and self-belief.

In terms of parent education background, generally can be concluded the lower the education of the parents, the lower achievement of the student. However, parent education background had no relationship with student' self-beliefs. For study support, compared to a student that has better study support (have room and other goods in-home), the student that have less study support perform lower in MATH, KNOW, and APPLY, but not in REASON. However, the relationship between home study support and SELLEFF was the opposite as in math achievement. The students who have either their room or other goods had positive SELFEFF, but for SELFCONCEPT, there was no relationship with home study support. Regarding this home study support, an empirical review by Filmer and Pritchett (1999) also found that the relationship between student's wealth gaps and academic performance had different pattern across countries. Therefore, the relationship between home resources in this study looked inconsistent too.

In terms of the number of books in a student's home, there is an inconsistent relationship between the outcomes. For example, having fewer books in comparison to no books had a negative association with MATH and KNOW. On the other hand, there was no relationship with APPLY, but having only 26-100 books had a positive correlation with REASON. However, the trend was mostly positive with this cognitive domain, indicating that the more book they have, the higher REASON score they have. This result is also similar to SELFCONCEPT and SELFEFF, which found that the number of books has a positive relationship with the student's self-belief (see Table 5-2 for score detail).

Moving to another socio-cultural background variable. The study uses language (how often students use the language of the TIMSS test (Indonesian language) in their daily life) as a proxy to measure ethnic differences. The result found that student who never uses the

language had a significant negative relationship with REASON, but not on other achievements. The trend was also similar to self-beliefs outcomes which the study found that also have a negative relationship with language use.

In terms of the student's psychological background variables, the study found that SELFCONCEPT has a significant positive relationship with MATH and all its cognitive domain. On the other hand, math achievement was also had a significant positive relationship with student self-beliefs. This result is in line with most of the studies in the area (Marsh, 1990a; Marsh & Martin, 2011; O'Mara et al., 2006; Pajares & Urdan, 2002; Parker et al., 2014). Therefore, this study also confirms the reciprocal relationship between achievement and self-beliefs (Huang, 2011; Marsh & O'Mara, 2008; O'Mara et al., 2006; Seaton et al., 2014).

Regarding the overall goodness of fit of the model, the percentage of total variance explained in Model 1 is very low, describing only 5% to 6% of the total variance in MATH, KNOW, and APPLY (see Table 5-2). For REASON and SELFCONCEPT, the total variance explained was 4% and only 1% for SELFEFF. This result suggests that considering student's characteristics factors in the model was hardly adequate to predict students' learning outcomes.

5.3.2 Teacher characteristics

There were four teacher-level variables included in the Model: gender, age, teacher's education background, and teacher's experience in teaching. Table 5-3 describes math achievement and its cognitive domain variability attributed to school and classroom after taking into account the teacher's background variables. The relationship between math achievement and all teacher's characteristics was statistically significant, except for gender. In terms of the self-beliefs outcome, however, none of the teacher variables had a relationship with student's self-belief.

In previous studies, the relationship between teacher's gender and student achievement was inconsistent. For example, teacher gender was found as a non-significant teacher variable

in Antecol et al. (2015) study who examined the effect of teacher gender on primary student's achievement using a randomised experiment. In other contexts, like Pakistan, teacher gender had a strong relationship with the student's achievement (Warwick & Jatoi, 1994).

The teacher's age was found to have a significant negative relationship with MATH, KNOW, APPLY, and REASON. The older the age of the teacher, the lower the student achievement in terms of their mathematic score and its cognitive domains, namely math knowing, applying, and reasoning. In contrast, a positive significant relationship was found for the teaching experiences variable. The teacher who had less experience tend to have lower student achievement. Furthermore, a similar result was also found on the positive significant relationship between teacher's educational background and student achievement. The students that were taught by a teacher who had no formal education beyond upper-secondary tended to have lower math scores. In contrast, the higher the teacher's educational background, and the more they have a specialty in math, were related to higher student achievement. Thus, the trends depicted that students' achievement in math was associated with several teachers' demographic composition, namely teachers' age, experience, educational background, and specialty in math.

Whereas the teacher's age was found to have a negative significant relationship with mathematics scores and its cognitive domain (KNOW, APPLY, REASON), teacher's experience, educational background, and speciality in math were found to have positive significant relationships. These results are supported by most of the research in the area which also found that teacher's experience and qualification positively influence student achievement (Croninger et al., 2007; Darling-Hammond, 2000a, 2000b). Surprisingly, the older teacher was associated negatively to student achievement. As age is not necessarily similar to experience, the older teachers might be less adapted to current teaching methods, or the wide generation gap hinders the interaction between teacher and students. However, further exploration is needed to justify how teacher age affects student achievement.

The variability of math score and its cognitive domain between school were slightly dropped, from 34% to 36% in Model 0 to 29% to 29% to 31%, in average, there was a 5% dropped in MATH, KNOW, and APPLY. For REASON, the variability between schools also dropped from 26% to 22%. In terms of classroom variability, the trend remained similar to Model 0. Unlike with achievement outcomes, SELFCONCEPT, and SELFEFF the variability between school and classroom was almost identical with Model 0.

Table 5-3: Model 2 teacher's backgrounds as explanatory variables

Response	**MATH**		**KNOW**		**APPLY**		**REASON**		**CONCEPT**		**EFFICACY**	
Fixed part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Cons	422.48	22.48	422.98	25.96	428.20	21.04	416.44	19.38	10.35	1.55	8.47	2.32
Gender												
Male	-2.02	8.26	-1.82	9.60	-3.06	8.32	-1.09	7.35	0.06	0.49	0.48	0.59
Age												
25-29	-24.15	17.05	-31.37	19.98	-28.89	16.18	-21.27	15.91	-1.36	1.26	0.24	1.86
30-39	-30.59	18.58	-38.76	21.25	-38.28	17.70	-28.82	17.02	-1.00	1.32	0.04	2.05
40-49	-13.44	20.33	-20.22	23.05	-19.83	18.82	-10.51	17.87	-0.65	1.43	0.83	2.27
50-59	-8.62	24.93	-13.12	29.06	-14.43	23.68	-5.70	21.91	-0.51	1.66	0.78	2.47
60 or more	-84.26	25.20	-102.83	29.38	-79.14	22.42	-77.92	23.20	0.66	1.57	3.04	2.40
Experiences												
At least 10 but less than 20 years	5.20	13.58	4.80	15.54	6.21	13.32	8.34	11.89	0.55	0.78	0.62	0.88
At least 5 but less than 10 years	-10.39	16.16	-10.35	18.68	-11.24	16.01	-6.93	14.32	1.79	1.00	2.21	1.21
Less than 5 years	-42.97	18.11	-47.70	20.66	-45.65	17.04	-33.53	15.30	1.47	1.19	2.75	1.61
All other majors	8.20	11.41	6.54	13.09	6.14	11.23	4.87	11.63	0.68	0.77	1.07	0.92
Education												
No formal education beyond upper-secondary	-60.97	16.77	-69.40	19.28	-64.85	15.15	-47.09	15.06	-0.58	1.27	1.89	2.24
Random Part												
school variance	1962.41	362.45	2577.66	476.51	1964.68	385.22	1538.62	289.43	4.09	1.57	6.86	1.83
class variance	565.16	236.90	747.48	284.13	584.22	245.54	378.40	167.10	3.49	1.42	4.79	1.59
Student variance	3906.80	102.24	5002.19	183.77	4196.25	112.56	5024.50	142.10	82.20	1.99	65.92	1.60
VPCschool	0.30		0.31		0.29		0.22		0.05		0.09	
VPCclassroom	0.09		0.09		0.09		0.05		0.04		0.06	
VPCstudent	0.61		0.60		0.62		0.72		0.92		0.85	

Deviance	64898.14	66327.46	65296.02	66263.19	42233.70	41042.46
School variance explained	19%	19%	21%	19%	9%	12%
Class variance explained	6%	-3%	-3%	3%	2%	0%
Student variance explained	0%	0%	0%	0%	0%	0%
Total Variance explained	7%	7%	7%	5%	0%	1%

Note:
Bold coefficient = $p < .05$Teacher gender is a dichotomous variable with Female as the reference group
Teacher age is categorical variables with age less than 25 as the reference group
Teacher experience is a categorical variable with more than 20 years' experience as the reference group
Teacher education background is a categorical variable with a major in math as the reference group.

Concerning the overall goodness of fit of the model, the percentage of total variance explained of Model 2 was poor but still slightly higher compared to Model 1, describing roughly 5% to 7% of the total variance in MATH, KNOW, APPLY, and REASON (see Table 5-3). For self-beliefs outcomes, the total variance explained become lower in SELFCONCEPT, and similar to Model 1 for SELFEFF. The total variance explained after including the teacher's background variables was still low, suggesting that considering the teacher's characteristics factors are not satisfactory to predict students' learning outcomes.

5.3.3 School characteristics

Model 3 (Table 5-4) was employed to examine school-level variables that explain math achievement and its cognitive domains, self-beliefs variance at the school and classroom level. School's social and economic background (aggregated from student's SES), school location, school size, and type of school (general private school, general public school, private *madrasah*, and public *madrasah*) were included in the model.

Table 5-4 showed that as expected schools with higher average SES tend to have significantly higher in MATH, KNOW, APPLY, and REASON. Next, regarding school location, students who attend schools in the small town achieved significantly lower than their counterparts in remote rural. Moving on to school types, also as expected, the student who attended private schools, both *madrasahs*, and the general school had significantly lower achievement.

Unlike the cognitive achievement outcomes, SELFEFF had a significant negative relationship with the school's average SES but had no relationship with SELFCONCEPT. For school location, a student in suburban had a significant positive relationship with their self-beliefs. Concerning school types, self-beliefs had no significant relationship SELFCONCEPT and SELFEFF.

Table 5-4: Model 3 school's backgrounds as explanatory variables

Response	**MATH**		**KNOW**		**APPLY**		**REASON**		**SELFCONCEPT**		**SELFEFF**	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Cons	414.20	8.80	410.93	10.50	411.25	9.13	409.63	8.03	9.21	0.47	8.21	0.73
Mean SES	32.06	5.59	37.04	6.57	31.51	5.80	26.90	5.07	-0.11	0.37	-0.86	0.41
Size	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Location												
Suburban	0.46	10.32	-0.79	11.90	3.36	10.32	4.35	9.19	1.37	0.57	1.70	0.76
Medium size city	-6.66	13.53	-6.28	15.39	-5.50	13.70	-2.51	11.89	1.61	0.90	2.20	1.01
Small town	-24.02	11.92	-26.84	13.57	-21.33	12.18	-17.95	10.81	1.39	0.77	3.06	0.98
Remote rural	-21.41	29.68	-31.71	34.02	-23.85	31.48	-17.25	27.81	2.05	1.49	3.21	2.74
Types												
Public *madrasah*	-8.10	12.69	-10.71	13.56	-10.16	12.21	-9.72	10.78	-0.64	1.11	0.11	1.19
Private general	-20.79	8.21	-23.74	9.53	-20.33	8.46	-20.90	7.57	-1.20	0.72	-0.73	0.75
Private madarasah	-34.05	12.19	-38.98	13.91	-37.44	12.79	-32.59	11.71	0.33	0.83	1.65	0.96
Random Part												
school variance	1222.16	300.82	1499.38	394.16	1235.95	333.61	956.72	262.58	3.09	1.70	4.03	1.66
class variance	659.29	242.43	869.44	292.83	678.74	257.55	469.18	182.45	3.88	1.56	6.10	1.65
Student variance	3907.00	102.21	5002.60	183.76	4196.69	112.53	5024.89	142.12	82.21	2.00	65.94	1.60
VPCschool	0.21		0.20		0.20		0.15		0.03		0.05	
VPCclassroom	0.11		0.12		0.11		0.07		0.04		0.08	
VPCstudent	0.67		0.68		0.69		0.78		0.92		0.87	
Deviance	64856.92		66279.78		65256.39		66224.07		42225.30		41027.01	
School variance explained	49%		53%		51%		50%		31%		48%	
Class variance explained	-10%		-19%		-19%		-21%		-10%		-27%	
Student variance explained	0%		0%		0%		0%		0%		0%	
Total Variance explained	16%		17%		16%		12%		1%		3%	

Bold p < .05
Location is a categorical variable with Urban as the reference group
Type is categorical variables with General public school as the reference group

Table 5-5: Model 4 student, teacher, and school's backgrounds as explanatory variables

Response	MATH		KNOW		APPLY		REASON		SELFCONCEPT		SELEFF	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Cons	448.93	23.76	448.29	26.15	449.68	21.88	433.92	20.26	9.60	1.71	5.99	2.42
Student level												
Gender												
Boy	-4.63	2.92	-8.05	3.27	-2.94	2.74	-1.23	2.56	0.34	0.28	0.67	0.28
Parents' education												
Post-secondary but not university	-9.75	6.14	-8.86	8.12	-8.51	5.56	-2.37	6.25	0.25	0.62	0.71	0.52
Upper secondary	-8.92	4.03	-8.40	4.51	-7.74	4.26	-11.25	5.35	0.50	0.42	0.50	0.40
Lower secondary	-12.43	3.57	-12.83	5.11	-11.48	4.06	-8.93	5.60	0.79	0.49	0.48	0.45
Some primary, lower secondary or no school	-11.18	5.00	-11.55	4.29	-16.00	4.35	-1.33	4.01	0.09	0.50	0.38	0.47
Home study support												
Either own room or electronic goods	-8.13	3.92	-5.39	4.11	-5.69	3.54	-6.06	3.44	0.08	0.32	0.77	0.34
Both own room and electronics goods	-8.13	4.82	-4.39	5.18	-5.82	4.33	-5.04	5.28	0.59	0.55	1.58	0.41
Number of books												
11-25 books	-6.12	2.68	-5.94	2.89	-6.10	3.39	-2.12	4.31	0.50	0.32	0.96	0.29
26-100 books	2.70	2.71	5.76	4.86	-2.20	4.54	8.37	3.91	0.67	0.40	0.92	0.40
101-200 books	-5.72	6.95	-11.77	7.99	-9.51	8.80	8.26	9.23	2.02	0.97	2.80	0.78
More than 200	-2.93	11.53	9.99	11.80	-25.22	12.75	3.21	12.64	3.62	0.91	3.92	1.06
Language												

Sometimes	-0.50	3.55	-0.10	4.15	0.56	3.10	-0.55	3.66	-1.06	0.30	-0.86	0.25
Never	-5.21	4.77	-4.61	5.68	2.59	5.37	-20.36	6.82	-2.68	0.51	-2.47	0.52
Self-Beliefs												
SELFCONCEPT	1.40	0.17	1.68	0.19	1.27	0.19	1.00	0.21	NA		NA	
SELFEFF	0.28	0.15	0.19	0.21	0.31	0.23	0.73	0.20	NA		NA	
Achievement												
MATH1	NA		NA		NA		NA		0.03	0.00	0.01	0.00
Teacher level												
Age												
25-29	-22.47	17.36	-28.31	19.82	-24.94	17.03	-21.93	16.57	-2.41	1.33	-0.61	1.86
30-39	-27.58	17.03	-34.18	19.40	-33.81	16.30	-27.54	16.05	-1.95	1.33	-0.88	1.91
40-49	-17.39	19.22	-23.77	21.38	-22.64	17.80	-16.40	16.98	-1.61	1.43	-0.01	2.08
50-59	-17.26	21.99	-21.04	25.23	-22.18	20.97	-13.68	19.63	-1.62	1.68	0.08	2.26
60 or more	-37.42	24.64	-44.45	28.60	-33.94	22.63	-35.84	23.38	0.81	1.83	2.53	2.24
Experience												
At least 10 but less than 20 years	10.68	11.42	11.88	12.75	11.94	11.40	13.00	9.93	0.27	0.77	0.19	0.86
At least 5 but less than 10 years	0.22	14.71	3.97	16.69	-0.57	14.69	2.94	12.91	1.72	0.97	1.97	1.12
Less than 5 years	-22.44	17.01	-21.40	19.45	-24.46	16.09	-16.93	14.18	0.94	1.21	1.87	1.46
Education background												
All other majors	14.80	10.28	15.21	10.92	13.81	10.21	10.94	10.18	0.65	0.79	1.21	0.90
No formal education beyond upper-secondary	-26.13	12.58	-26.23	14.10	-29.53	13.38	-19.93	13.18	-1.08	1.16	1.14	1.63
School-level												
Mean SES	27.94	6.18	32.17	6.97	27.39	6.19	23.63	5.50	-0.56	0.42	-1.28	0.45
Location												
Suburban	4.34	10.76	3.35	12.05	5.84	10.82	7.02	9.62	2.01	0.67	2.04	0.88
Medium size city	-2.51	13.55	-1.36	14.89	-2.27	13.66	0.86	11.67	2.36	0.97	2.62	1.07

Small town	-17.90	13.03	-21.28	14.37	-16.64	13.12	-13.75	11.48	1.46	0.84	2.98	1.16
Remote rural	-12.97	29.77	-23.56	34.23	-15.74	31.25	-11.97	28.05	2.32	1.44	2.88	2.47
Types												
Public *madrasah*	0.13	14.10	-3.08	15.42	-0.46	13.78	-1.32	12.54	-1.00	1.12	-0.19	1.22
Private general	-13.64	9.49	-17.56	10.80	-13.24	9.67	-14.83	8.89	-1.89	0.83	-1.68	0.83
Private *madrasah*	-30.45	12.14	-35.36	13.96	-33.02	12.84	-29.92	11.82	0.08	0.83	1.21	0.94
Random Part												
school variance	1081.26	295.20	1330.98	366.08	1096.25	301.21	890.43	239.92	2.68	1.77	4.10	1.45
class variance	663.39	245.27	872.44	291.68	670.04	254.73	435.90	181.04	4.12	1.54	5.49	1.38
Student variance	3675.29	98.69	4684.51	177.32	4001.77	121.98	4776.43	154.33	77.92	1.80	64.16	1.55
VPCschool	0.20		0.19		0.19		0.15		0.03		0.06	
VPCclassroom	0.12		0.13		0.12		0.07		0.05		0.07	
VPCstudent	0.68		0.68		0.69		0.78		0.92		0.87	
Deviance	64494.98		65898.86		64975.97		65926.46		41918.28		40864.63	
School variance explained	55%		58%		56%		53%		40%		47%	
Class variance explained	-11%		-20%		-18%		-12%		-16%		-14%	
Student variance explained	6%		6%		5%		5%		5%		3%	
Total Variance explained	22%		23%		21%		17%		6%		6%	

The variability of math score and its cognitive domain between school was considerably dropped up to 16% compared to Model 0 in MATH, KNOW, and APPLY. For REASON, the variability percentage attributable to differences between schools also dropped from 26% to 15%. In terms of classroom variability, the trend remained similar to Model 0, with slightly increased up to 2%. Consistent with previous research which mostly found that when the school context is taken into account, the apparent school effect is reduced (Muijs & Reynolds, 2003; Opdenakker & Van Damme, 2000; Opdenakker et al., 2002). However, unlike achievement outcomes, SELFCONCEPT and SELFEFF variability between school and classroom was almost identical with Model 0, with 2% in SELFEFF.

By taking into account the school context in Model 3, the overall "goodness of fit" of the model was significantly improved compared with Model 1 and Model 2, explaining 16% to 17% of the total variance in MATH, KNOW, and APPLY. For REASON the total variance explained was slightly lower (12%) compared to other cognitive domains. However, for SELFCONCEPT and SELFEFF, the change was minimal. This goodness of fit was considerably better to predict students' learning outcomes. It indicated that school factors are sounder in explaining student achievement compared to student's and teacher's background variables.

To explore the range and extent of school performance in Indonesian lower secondary schools in math and self-beliefs after controlling for student, teacher, and school characteristics, Model 4 (Table 5-5) was performed. For math and its domains, the variance between schools looked similar to Model 3 but slightly dropped up to 1%. However, for SELFEFF, there was a 1% increase from the previous model. However, there were some changes worth mentioning. For example, teacher characteristics, both experiences, and education background were found to have no statistically significant relationship with math and its domains. But the teacher's age was found to have a negative significant relationship with APPLY only. For Self-beliefs

outcomes, as in Model 2, teacher's characteristics were found to have significant relationship with these non-academic outcomes.

For student characteristics variables, the effect remained the same as Model 1. Moreover, in the school context variables, the effect was found to be similar to Model 3, as well. However, for school types, a change was found. Where the private school in Model 4 now had no significant relationship with all the outcomes, the direction was similar to Model 3. In Model 4, only private *madrasah* had a significant negative relationship with math achievement and its cognitive domains. It indicated that, after including student, teacher, and school background variables, a student who attended private *madrasah* was found to have a lower academic achievement compared to other students in other school types (public general, public *madrasah*, and private general). On the other hand, for SELFCONCEPT and SELFEFF, student, teacher, and school-level variables had a relatively similar significant effect as Model 1, Model 2, and Model 3.

Including all student, teacher, and school-level variables in the model, the overall "goodness of fit" was considerably improved compared to Model 3, explaining 21% to 23% of the total school variance in MATH, KNOW, and APPLY, improved about 5-6%. For REASON, the total variance explained was also higher in Model 4. There was a 5% variance increase compared to Model 3, from 12% to 17%. In terms of school classroom variability, compared to Model 0, 1, and 2, Model 4 had a remarkable improvement in school variability, but slightly higher compared to Model 3. On the other hand, classroom variability was not improved significantly for all outcomes.

The next section explains the main factors examined in the study, school climate factors, and then combining all the background variables at the student, teacher, and school levels.

5.4 RQ1.3: School climate factors that significantly explain the variance between school and classroom performance

RQ1.3 asked, "What are the school climate factors that significantly explain the variance between school and classroom performance among Indonesian Year 8 students in math and self-beliefs before and after adjusting the characteristics of the student, teacher, and school?" To answer RQ1.3, Model 5 and Model 6 was conducted to compare the school climate factors that significantly explain the variance between school and classroom performance before and after controlling the characteristics of the student, teacher, and school.

5.4.1 Significant school climate factors before adjusting the characteristics of the student, teacher, and school

This section discusses the results of Model 5. Model 5 included all school and classroom climate variables taken from different levels (student and teacher questionnaire) that aggregated at the school and classroom level. This combination was taken into account because as clearly explained in Chapter 3. The debate is mainly concerning two main viewpoints. The first one claimed that the school climate is an individual-level variable (James, 1982; Miller & Fredericks, 1990). The second viewpoint claimed that all school members experienced a similar climate through their shared interaction within the same setting (Van Horn, 2003).

Model 5 (Table 5-6), having accounted for all school climate factors only, 23% to 25% of the variance in MATH, KNOW, APPLY was attributable to the difference between schools. On the other hand, for REASON outcome, the variance between schools was considerably lower, approximately 17%. This variance explained was significantly improved compared to Model 0. The improvement was about 11% for academic outcomes. It indicated that school climate factors have a considerable effect on student's achievement. However, the classroom differences were almost the same with Model 0, but slightly lower, about 1% dropped. As

mentioned before this was because the classroom sample variability was relatively small, so may not detect variability accurately.

Table 5-6: Model 5 school climate factors

Response	**MATH**		**KNOW**		**APPLY**		**REASON**		**SELFCONCEPT**		**SELFEFF**	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Cons	398.22	3.67	391.60	4.18	397.13	3.90	397.70	3.51	10.30	0.19	10.19	0.20
School climate												
Student connected with school	0.35	0.14	0.44	0.22	0.22	0.25	0.51	0.23	0.04	0.02	0.07	0.02
Classroom climate												
Student engagement with math lesson (teaching-learning)	0.04	0.21	0.13	0.24	0.27	0.22	-0.19	0.31	-0.23	0.03	-0.42	0.02
Student safety	-0.67	0.29	-0.68	0.29	-0.78	0.28	-1.02	0.28	0.15	0.03	0.02	0.03
Teacher safety	0.72	1.44	0.84	1.67	0.70	1.51	0.72	1.46	0.12	0.03	0.11	0.05
Teacher emphasis on academic success	2.76	2.42	3.05	2.87	3.22	2.51	2.30	2.34	0.01	0.07	-0.02	0.10
Teacher-teacher interaction	0.48	2.31	1.66	2.58	0.51	2.44	-0.22	2.31	-0.03	0.09	-0.28	0.11
Teacher connected with school	-1.16	1.55	-0.88	1.65	-1.01	1.48	-1.32	1.46	-0.10	0.05	-0.07	0.07
Teacher working condition	0.32	2.67	-0.53	3.08	0.00	2.75	0.26	2.63	-0.10	0.09	-0.18	0.10
Teacher: instruction to engage student	-1.16	2.39	-2.14	2.64	-1.58	2.37	-0.49	2.15	-0.06	0.05	0.01	0.06
Teacher: classroom disturbance	-0.15	3.13	-0.43	3.53	-0.84	3.31	1.15	3.04	-0.01	0.08	0.15	0.09
Teacher: confident in teaching	-0.01	2.58	0.34	3.15	0.74	2.70	-1.01	2.66	0.35	0.08	0.32	0.10
School: emphasis on academic success)	2.51	0.91	2.91	1.05	2.42	0.91	2.14	0.83	0.01	0.05	-0.06	0.05

School: discipline	2.25	0.97	2.63	1.11	2.13	0.99	2.09	0.86	0.06	0.06	0.06	0.06
School: technology resources	0.15	0.62	0.28	0.71	0.19	0.61	0.20	0.54	-0.01	0.03	0.00	0.03
School: safety	-0.59	0.67	-0.81	0.77	-0.56	0.69	-0.49	0.61	-0.02	0.05	-0.03	0.05
School: general resources	-2.65	1.08	-3.08	1.28	-2.79	1.10	-2.12	0.97	0.08	0.06	0.15	0.07
School leadership	-0.99	0.61	-1.05	0.69	-0.89	0.61	-0.82	0.55	-0.02	0.03	0.03	0.03
School mean: student connected with school	1.94	1.81	2.23	2.10	2.14	1.83	2.47	1.69	-0.08	0.10	0.00	0.10
School mean: student safety	-7.18	2.46	-7.93	2.79	-7.88	2.47	-6.25	2.26	0.50	0.14	0.42	0.17
School mean: teacher connected with school	0.91	0.67	0.96	0.78	0.81	0.69	0.83	0.62	-0.01	0.04	-0.03	0.04
School mean: teacher emphasis on academic success	0.30	0.74	0.18	0.84	-0.03	0.74	0.10	0.66	0.00	0.03	-0.01	0.04
School mean: teacher-teacher interaction	-0.61	0.70	-0.83	0.77	-0.41	0.69	-0.75	0.63	-0.03	0.03	0.04	0.04
School mean: teacher safety	-0.72	0.31	-0.86	0.37	-0.62	0.32	-0.71	0.29	-0.01	0.02	0.00	0.02
School mean: teacher working condition	-2.78	1.52	-2.92	1.73	-2.84	1.51	-2.10	1.39	-0.11	0.07	0.06	0.07
Class mean: teacher confident in teaching	0.57	0.94	0.59	1.05	0.31	0.96	0.33	0.83	-0.02	0.05	-0.07	0.05
Class mean: teacher engaging instruction	-0.13	0.55	-0.12	0.61	0.04	0.54	0.09	0.50	0.04	0.03	0.02	0.03
Class mean: classroom disturbance	-0.18	0.71	-0.40	0.77	-0.32	0.70	-0.39	0.68	0.01	0.04	0.03	0.05
Class mean: student engagement in math lesson	6.20	1.62	7.15	1.95	6.93	1.71	5.39	1.57	-0.78	0.08	-0.92	0.11
Random Part												

school variance	1399.20	279.65	1841.55	349.77	1392.65	275.56	1075.12	221.18	2.94	1.01	3.59	1.12
class variance	472.98	170.73	594.87	202.33	461.21	164.79	356.17	138.19	0.30	0.89	1.35	1.05
Student variance	3888.48	102.21	4980.31	179.38	4176.88	110.46	4979.77	139.62	79.56	1.99	58.59	1.44
VPCschool	0.24		0.25		0.23		0.17		0.04		0.06	
VPCclassroom	0.08		0.08		0.08		0.06		0.00		0.02	
VPCstudent	0.68		0.67		0.69		0.78		0.96		0.92	
Deviance	64818.49		66253.60		65219.88		66170.52		41947.56		40256.40	
School variance explained	42%		42%		44%		44%		34%		54%	
Class variance explained	21%		18%		19%		8%		92%		72%	
Student variance explained	0%		0%		0%		1%		3%		11%	
Total Variance explained	17%		17%		17%		12%		8%		19%	

For self-belief, the variance between schools was dropped about by 4%, from 10% in Model 0 to 6% in Model 5 for SELFEFF. For SELFCONCEPT, there was only 1% dropped from 5% to 4% For SELFCONCEPT. However, for classroom variability, the two self-beliefs outcomes were 4% dropped compared to Model 0.

The percentage of total variance explained (overall goodness of fit) of Model 5 in self-beliefs outcome was relatively higher than Model 1, 2, 3, and 4. A significant improvement was on SELFEFF which explains 19% of total its total variance explained. This result indicated that the school climate factor also had a significant effect on student self-beliefs outcomes.

In terms of school climate factors, not all the factors had a significant relationship with all learning outcomes. At the student level, student connection with school (like being in school) had a significant positive relationship with MATH, KNOW, REASON, and SELFEFF. Then, student engagement in math lesson only was found to have a significant negative relationship with both self-beliefs but not with math achievement. This result was dissimilar with previous research, for example, Fung et al. (2018) examined 295,416 15-year-old secondary school students from 34 using PISA 2012 data. They found that students who were more engaged in math lessons had higher levels of academic achievement. In the Indonesian context, the student perceived that they were engaged in school and classroom, but their achievement result was found to be low (see Figure 5.1).

For student safety, the result was unexpected because the finding showed a significant negative relationship between student safety and academic achievement. However, student safety had a significant positive relationship with SELFCONCEPT. Perhaps the potential reason is that the correlation was not linear.

At the teacher level, teacher-teacher interaction with their colleagues also had a significant negative relationship with SELFEFF. Next, teacher confidence in teaching math had a significant positive relationship with student's self-belief.

At the school level, school emphasising on academic success had a significant positive relationship with student achievement, but not with student self-beliefs. Then school's general resources unexpectedly had a significant negative relationship with student achievement but had a positive relationship with SELFEFF. Lastly, school leadership, unpredictably had no significant relationship with all learning outcomes.

Next, the climate factors aggregated at school and classroom level were described. School *mean* of connectedness had no significant relationship with any student learning outcomes. On the other hand, school mean of student safety also had a similar pattern as in non-aggregated factors, the relationship was negatively significant with math achievement and had a significant positive relationship with self-beliefs. For the teacher questionnaire, the school mean of student safety also had a significant negative relationship. Lastly, classroom mean of student engagement had a significant positive relationship with math achievement and all its cognitive domains but had a significant negative relationship with students' self-beliefs.

The next model (Model 6) included all significant school and classroom climate factors as well as all significant student, teacher, and school background variables.

5.4.2 The impact of school climate factors after adjusting the characteristics of the student, teacher, and school

To determine if there are any differences in math and its cognitive domains, SELFCONCEPT, and SELFEFF between schools and classrooms after including all significant school climate factors in Model 5 and adjusting for the student, teacher, and school characteristics in Model 4, Model 6 was estimated (Table 5-8).

Table 5-8 showed that variance attributable to differences between schools in student MATH, KNOW, and APPLY drop significantly to 18% – 19%. For REASON the variance between schools drops to 14%. However, for SELFCONCEPT and SELFEFF, the variance

was attributable to differences between schools was not significantly different compared to previous models (Model 3 and Model 4).

Model 6 showed the contribution of schools and classrooms climate to student learning is better than Model 4 (all student, teacher, and school background) and 5 (Schools and classrooms factors only). The percentage of total variance explained in Model 6 is higher than those two models, explaining approximately 25% - 26% of the total variance in MATH, KNOW, and APPLY; and 19% in REASON. For SELFCONCEPT and SELFEFF, the total variance explained considerably improved, explaining 13% and 22% respectively. That total variance explained is not significantly high but can prove the effectiveness of school and classroom climate on student achievement and self-belief. 25% total variance explained was acceptable since this model did not include one of the most significant predictors e.g., prior achievement (Timmermans & Thomas, 2014). For comparison, Muñoz-Chereau (2013) in her research found that by adding prior attainment, the goodness of fit improved dramatically compared to a model that only includes student background variables, explaining 63% compared to 16% respectively.

Regarding schools and classrooms variance explained, this model explained 57% to 60% school variance explained in math and its cognitive domains, 70% in SELFEFF, and 46% in SELFCONCEPT. Interestingly, SELFCONCEPT was found to have the highest classroom variance explained, approximately 65%.

For student's characteristics, gender was found to have no significant relationship with student academic achievement but have a significant relationship with SELFCONCEPT and SELFEFF, where the boys were found to have a higher score in self-belief compared to the girls. These results are consistent with previous research such as Hergovich et al. (2004). They found that girls' self-concept in math was lower compared to a boy. They added, girls' self-concept much depends on teachers' and parents' judgments, whereas boys are not.

For parents' education, the number of books, the result showed relatively similar to Model 1. Next, for teacher characteristics, the teacher's age was found to have a significant relationship with student achievement. However, the relation was not linear, because student tends to have lower achievement with younger (less than 40 years old) and with older teachers (more than 60 years old). This result indicated that teachers of average age (40-59) have experiences needed in managing their classrooms. Teachers under the '30s considerably have not enough experience, and they are relatively young, so they may not understand their students' behaviour better. Contrary, teachers between 40 to 59 have more experience.

For self-belief outcomes, teachers' age has no significant factor, but teachers' major was a matter. The student with a non-mathematics teacher tended to have higher SELFCONCEPT and SELFEFF.

Regarding school level characteristics or school context, average SES was found to have a significant relationship with student achievement but not on student self-beliefs. This result was consistent with other research mostly found that schools with a low-SES group of students are often under-resourced and can affect student attainment (Aikens & Barbarin, 2008; Timmermans & Thomas, 2014). Also, other research found that the socioeconomic composition of students had a significant effect on student achievement (Muijs, et al., 2010; Timmermans & Thomas, 2014).

5.5 The achievement gap between general school and *madrasah*

RQ1.4 asked: Why do students from the general school have higher achievement and self-beliefs than those from *madrasah* before and after controlling school climate and other factors? To answer RQ1.4, Model 4 (including only significant variables of the student, teacher, and school characteristics) and Model 6 (including all significant school climate factors and all significant variables of the student, teacher, and school characteristics) were compared. The results are presented in Table 5-7.

School types (general/*madrasah*), together with another school-level factor, school location, were found to have no significant relationship with all the student learning outcomes (academic and self-beliefs). This result was unexpected and interesting, because by taking into account schools and classroom climate factors, the effect of school types (particularly being private *madrasah*) became insignificant. This finding indicated that schools and classroom climate factors are essential in reducing the negative effect of being private schools regarding its type compared to Model 4 (where private *madrasah* had significantly lower academic achievement compared to other school types).

Compare to Model 5; this model showed that only five schools and classroom climate factors now had a statistically significant relationship with student learning outcomes. As expected, based on previous research, student engagement in a math lesson was significantly related to MATH, KNOW, and APPLY, but not with REASON. This result is in line with most school climate research that found the teaching-learning process is the most crucial factor in a school's climate that explains student achievement. In schools that prioritising on academic success, academic performance tends to be higher (see Opdenakker & Van Damme, 2005).

However, this factor (student engagement in maths lessons) had a negative relationship with SELFCONCEPT and SELFEFF. Indicating that if students perceived that all students in the classrooms engaged in a math lesson, their self-concept and self-efficacy tend to be lower. The possible explanation of this finding is because these two self-beliefs are sensitive to the comparison. If students perceived that other students are better than them, students tend to have lower self-beliefs.

Another key climate factor was safety including school discipline and safety. However, the direction of the relationship between these two safety factors was found to be different.

Table 5-7: Model 6 Final Model

Response	MATH		KNOW		APPLY		REASON		SELFCONCEPT		SELFEFF	
Fixed Part	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
cons	436.65	13.17	437.15	14.08	436.28	12.72	425.47	12.21	10.26	1.05	8.02	1.33
Student Level												
Gender												
Boy	-4.84	2.94	-8.30	3.25	-3.46	2.66	-1.27	2.52	0.50	0.26	0.95	0.24
Parents' education												
Post-secondary but not university	-9.58	6.19	-8.57	7.99	-8.15	5.55	-2.38	6.28	-0.04	0.60	0.19	0.49
Upper secondary	-8.88	4.05	-8.30	4.44	-7.59	4.25	-11.34	5.37	0.35	0.42	0.18	0.38
Lower secondary	-12.26	3.54	-12.58	5.03	-11.20	4.04	-8.82	5.64	0.61	0.49	0.18	0.45
Some primary, lower secondary or no school	-10.95	5.01	-11.25	4.23	-15.57	4.38	-1.17	4.01	-0.11	0.49	-0.03	0.46
Home study resources												
Either own room or electronic goods	-7.43	3.82	-4.61	4.16	-4.70	3.58	-5.29	3.46	-0.15	0.31	0.40	0.33
Both own room and electronic good	-7.72	4.69	-3.91	5.18	-5.14	4.20	-4.66	5.23	0.34	0.52	1.13	0.38
Number of books												
11-25 books	-6.16	2.66	-5.98	2.90	-6.13	3.38	-2.20	4.30	0.42	0.31	0.83	0.26
26-100 books	2.69	2.70	5.80	4.91	-2.00	4.58	8.14	3.92	0.43	0.39	0.46	0.37
101-200 books	-6.08	6.95	-12.21	8.03	-9.73	8.76	7.64	9.22	1.62	0.93	2.24	0.68
More than 200	-3.49	11.32	9.41	11.74	-25.73	12.62	2.41	12.59	3.32	0.96	3.26	1.04
Language												
Sometimes	-0.47	3.50	-0.17	4.13	0.48	3.10	-0.38	3.60	-0.92	0.29	-0.57	0.25
Never	-4.20	4.65	-3.62	5.60	3.58	5.46	-18.88	6.79	-2.51	0.51	-1.90	0.48
Self-beliefs												
SELFCONCEPT	1.45	0.17	1.74	0.18	1.34	0.19	1.06	0.21	NA		NA	

SELFEFF	0.36	0.16	0.29	0.23	0.45	0.25	0.75	0.23				
Achievement												
MATH1	NA		NA		NA		NA		0.03	0.00	0.01	0.00
Teacher Level												
Age												
25-29	-24.98	12.78	-30.92	14.67	-28.80	12.52	-22.91	13.15	-0.39	0.92	1.30	1.14
30-39	-22.40	10.44	-29.26	11.96	-28.11	10.38	-22.73	10.82	-0.43	0.83	0.20	1.04
40-49	-8.95	10.97	-16.86	11.82	-13.92	10.16	-8.56	10.41	-0.55	0.83	0.72	1.11
50-59	-9.35	13.76	-14.84	15.59	-14.35	13.33	-7.72	13.39	-0.79	1.11	0.61	1.25
60 or more	-30.92	18.20	-40.95	18.18	-23.20	15.91	-36.08	17.09	0.68	1.09	2.16	1.27
Education background												
All other majors	13.18	8.98	13.29	9.87	11.31	8.91	10.12	9.40	1.17	0.55	1.79	0.69
No formal education beyond upper-secondary	-22.78	14.14	-24.90	16.04	-26.55	16.58	-15.43	15.05	-1.13	0.76	1.43	0.91
School-level												
Mean SES	20.86	6.06	23.92	6.79	19.55	5.94	17.60	5.49	-0.12	0.38	-0.55	0.39
Location												
Suburban	9.69	9.44	9.85	10.96	12.41	9.78	10.64	8.92	0.59	0.63	-0.14	0.67
Medium size city	0.68	12.84	2.66	14.59	2.78	13.07	1.88	11.65	0.72	0.79	0.28	0.75
Small town	-8.61	11.03	-10.14	12.41	-6.04	11.28	-6.17	9.96	0.34	0.74	1.13	0.82
Remote rural	0.05	28.44	-8.51	32.76	-0.15	30.03	-1.05	27.50	0.89	1.12	0.55	1.86
Types												
Public *madrasah*	-6.40	15.35	-10.09	16.73	-8.86	14.82	-6.08	14.06	-0.42	0.92	0.85	0.78
Private general	-11.50	8.06	-15.38	9.23	-12.03	8.22	-12.33	7.48	-0.76	0.55	-0.99	0.58
Private madarasah	-22.11	12.52	-27.52	14.11	-24.80	12.99	-22.14	12.02	-0.12	0.67	0.54	0.72
School climate												
Student connected with school	0.23	0.14	0.28	0.21	0.12	0.26	0.37	0.22	0.03	0.02	0.07	0.02
Student safety	-0.89	0.28	-0.94	0.29	-1.01	0.28	-1.13	0.27	0.17	0.03	0.03	0.03

Teacher safety	0.43	0.98	0.53	1.13	0.51	0.99	0.35	1.00	0.04	0.03	0.05	0.03
Teacher-teacher interaction	-0.76	1.18	-0.92	1.48	-1.02	1.44	-0.68	1.27	-0.03	0.09	-0.14	0.05
Teacher confidence in teaching math	-0.87	1.38	-1.31	1.42	-0.72	1.51	-1.38	1.32	0.12	0.08	0.15	0.06
School emphasis on academic success	0.95	0.80	1.18	0.91	1.05	0.81	0.73	0.74	0.02	0.04	-0.02	0.05
School discipline	1.39	0.55	1.49	0.63	1.31	0.55	1.46	0.52	0.03	0.03	0.04	0.04
School's general resources	-1.50	0.98	-1.63	1.12	-1.65	1.00	-1.21	0.91	0.09	0.05	0.16	0.06
Student safe (school)	-5.09	2.36	-5.71	2.66	-6.16	2.35	-3.93	2.15	0.46	0.14	0.39	0.16
Teacher safe (school)	-0.31	0.33	-0.40	0.39	-0.33	0.35	-0.37	0.32	-0.01	0.02	-0.01	0.02
Classroom climate												
Student engagement in math lesson (class)	4.81	1.32	5.51	1.51	5.30	1.41	4.08	1.20	-0.75	0.07	-0.95	0.09
Student engagement in math lesson	0.51	0.21	0.67	0.26	0.74	0.20	0.39	0.32	-0.23	0.03	-0.41	0.02
Random Part												
school variance	1006.21	275.83	1265.93	335.99	1000.48	264.76	824.37	207.48	2.41	1.02	2.36	0.92
class variance	518.05	207.09	653.69	246.15	501.00	196.11	334.85	150.31	1.06	0.87	1.90	0.90
Student variance	3649.19	99.20	4650.08	168.44	3962.46	117.48	4737.83	149.25	75.33	1.84	57.15	1.40
VPCschool	0.19		0.19		0.18		0.14		0.03		0.04	
VPCclassroom	0.10		0.10		0.09		0.06		0.01		0.03	
VPCstudent	0.71		0.71		0.73		0.80		0.96		0.93	
Deviance	64431.66		65833.41		64892.92		65858.81		41646.41		40100.52	
School variance explained	58%		60%		60%		57%		46%		70%	
Class variance explained	14%		10%		12%		14%		70%		61%	
Student variance explained	6%		7%		6%		6%		8%		13%	
Total Variance explained	25%		26%		25%		19%		13%		22%	

School discipline had a significant positive relationship with students' achievement. In the contrary, school safety had a significant negative relationship with student achievement. This result is inconsistent with previous research that mostly found that student perception of safety had a significant positive relationship (see Thapa et al., 2013). However, this factor was found to be associated differently with both self-beliefs outcomes. The relationship between self-beliefs and safety had a significant positive relationship.

Moreover, the schools' general resources had a significant positive relationship with SELFEFF. This result is consistent with most of the school climate research that suggested the importance of the physical learning environment (see Anderson, 1982; Thapa et al., 2013; UNESCO-UIS, 2012). This result is also consistent with another researcher who mostly but not fully agreed that the adequacy of school facilities helps student learning (Shernoff, 2013).

5. 2. Chapter summary

In this chapter, school climate factors explaining schools and classrooms variance in math and its cognitive domain, and student self-beliefs were explored. In comparison to the null model, the percentage of total variance explained in the final model was 25% - 26% for MATH, KNOW, and APPLY; and only 19% for REASON. This total variance explained was moderate since there was no previous attainment included in the model. For SELFCONCEPT and SELFEFF, the total variance also explained considerably low, explaining only 13% and 22% respectively. However, this low goodness of fit was acceptable since this model did not include one of the most significant predictors, prior achievement (Newhouse & Beegle, 2006). There is no prior achievement provided in TIMSS data. Correspondingly, this finding did not measure effectiveness and therefore, should be interpreted cautiously.

Regarding school climate factors, in the final model, there were only two school climate factors that found to have a significant relationship with math performance, and five factors that related to self-beliefs. The two factors related to math achievement are: (1) *safety* as

measured by students not being bullied, and school discipline, and (2) *teaching and learning* as measured by student engagement in a math lesson. On the other hand, five factors that were related to the self-beliefs outcome are: (1) safety: student safety (not being bullied) and (2) teacher safety. Next, (3) relationship as measured by teacher-teacher interaction; (4) teaching and learning as measured by student engagement in a math lesson; and (5) institutional environment as measured by general school resources and students connected with the schools.

Concerning safety factors, the student who reported feeling safe in schools tends to have lower academic achievement. However, this factor was found to have a significant positive relationship with student self-concept. Students who felt safe at schools tend to have positive self-efficacy.

This result was inconsistent with previous research in school climate areas that mostly found that safety is one of the school climate factors that plays an essential role in student achievement. The possible explanation for this is because in the Indonesian context what consider as bullying may be different, or a low understanding of what can be regarded as bullying (Widayanti & Siswati, 2009). For example, Lai et al. (2008) found that the most popular form of bullying in the Asia Pacific region, including Indonesian secondary schools is of ‘students being made fun of or being called names. In the Indonesian context, for some reason that is not considered bullying (Widayanti & Siswati, 2009).

For teaching and learning factors, student engagement in a math lesson, which was significantly related to MATH, KNOW, and APPLY, but not with REASON. These results are in line with most school climate research that found that the teaching-learning process is the most crucial factor in a school's climate that, in turn, explaining student achievement. However, this factor had a negative relationship with SELFCONCEPT and SELFEFF, indicating that if students perceived that all students in the classrooms engaged in a math lesson, their self-concept and self-efficacy tend to be lower. The possible explanation is because these

two self-beliefs are sensitive to the comparison. If the student perceived that other students are better than themselves, the student tends to have lower self-beliefs.

Another factor, as a proxy of the institutional, physical environment, school resources reported by headteacher have a significant positive relationship only with student self-efficacy, indicating sufficient school resources lead to resulting in more positive self-efficacy. These results are consistent with most schools' climate research that suggested the importance of a physical learning environment (see Anderson, 1982; Hoy et al., 2006; Thapa et al., 2013; UNESCO-UIS, 2012).

Relating to the differential effect on different learning outcomes, school climate factors are more influential on academic achievement compare to self-belief outcomes. Regarding the difference between general and *madrasah* and other school types, this study found that by taking into account schools and classroom climate factors, the effect of school types, mainly being private *madrasah* became insignificant. This finding indicates that school and classroom climate factors may be essential in reducing the negative effect of being a private *madrasah* (Model 6). Compared to Model 4 (before accounting for schools and classrooms climate factors), the result found that private *madrasah's* student tends to have lower academic achievement.

Lastly, another result emerged in this research. By including different mathematic cognitive domains on the model, this research found that secondary schools in Indonesia are mainly varied, it showed more significant differences in student performance was stressed on KNOW and APPLY domains, rather than concerning higher order thinking as measured in REASON. This is because, in the Indonesian context, the teacher typically may be forced to teach a lot of subjects as in the curriculum, since these will be tested in the exam. Therefore, the teacher tends to teach questions that will possibly appear in the exam (teaching to the test rather than teaching for understanding). From Model 0 to the final model, REASON had fewer

differences between schools compare to other domains. The reasoning domain is associated with having the competence of logical and systematic thinking, with solving non-routine problems (Mullis et al., 2009).

In the following chapter, the findings depicted from qualitative data analysis are presented and discussed to explore school members' views of school climate concerning student learning outcomes in the Indonesian secondary schools. It explores further explanation to describe the differences in schools and classroom climate practices between four selected schools. It provides a more detailed explanation that offers a much better understanding of the variations in the school climate in Indonesia.

Chapter 6: Qualitative findings

6.1 Introduction to the chapter

The researcher has shown the significant effect of school climate in explaining students learning outcomes in Chapter 5. The present Chapter 6 aims to present the results of the qualitative phase of the research. Specifically, the chapter is aimed at answering the fourth research objective and RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)?

- RQ2.1: (a) What are the similarities and differences in school climate between high and low-performing schools? (b) and between religious/non-religious school settings?
- RQ2.2: Are there new factors that can be obtained from a qualitative inquiry that are relevant to highlight differences between the high and low performing schools?

As in the previous chapter, the researcher will answer the research questions one by one in sequence.

6.2 Stakeholders' perception of school climate

RQ2.1 addresses a question on how school stakeholders perceived the climate of their school. Classification have been made as follow: (1) Similarities across all school, and (2) differences between high and low performing schools, and (3) differences between religious/non-religious (*madrasah* and general) school.

6.2.1 Similarities across all schools

All schools showed three out of the five school climate factors suggested by Thapa et al. (2013). Namely, relationship, safety, and teaching and learning process factors. However,

by taking account the DMEE evaluative dimensions, similarities concerning the relationship factor in all four schools was observed only in terms of the frequency, stage, and differentiation dimensions. The security factor was similar only in terms of the frequency, focus, stage, and differentiation dimensions. Finally, the teaching and learning process factor was similar only in terms of the frequency dimension. This section describes each school climate factor with its DMEE evaluative dimensions, that appeared to be consistent across all schools.

6.2.1.1 Safety

As explained in Chapter 3, the safety factor consists of several forms: physical safety, emotional safety, school rules and norms, and actions taken for violating these rules and norms (Thapa et al., 2013, Cohen et al., 2009). The present study found that the four schools had relatively similar safety features, as explained below.

6.2.1.1.1 Safety: frequency

In general, stakeholders in all schools felt that students felt emotionally safe at their school. They reported that disturbances do take place at times, but nothing alarming. Cases of disturbance that had taken place were nothing near to bullying or things that endanger the emotional well-being of students.

Students usually respect each other. They are not annoyed or upset. They usually joke and laugh together (Guru, Madrasah A).

No bullying happens in this school. Threat and intimidation between one student against another rarely happen, because we have strict rules (Guru, Madrasah B)

However, it should be noted that in Indonesia and other Asian countries, teachers' understanding of what can be considered bullying is limited (Widayanti & Siswati, 2009). For example, making fun of or giving names is not usually considered a form of bullying (Lai et

al., 2008). As such, the teachers and student's perception concerning bullying may be bias in a sense that they may have failed to acknowledge certain students' behaviours as bullying.

In terms of rules and norms, all four schools also have written rules and norms to promote the safety of the students. The frequency dimension of this safety sub-factor discusses the availability of documents that regulate rules and norms in schools. In the four schools being studied these documents were given to students before or on the first day of school. The rules and norms are socialised not only among students, but also the parents.

It was given on the first day of school. Also, at the time of student orientation, prior to the formal start of school, they were informed about school rules and norms (Teacher, School D)

We have clear rules in both in the school and classroom. The rules are presented in the new school term. So, parents and their children already know the rules from the beginning. If a problem happens during the school process, the parent is immediately reported (Teacher, Madrasah B)

6.2.1.1.2 Safety: focus

From the students' point of view, rules are something that they should comply with. They are especially attentive to the penalties for rule breaking. For example, students are required to wear school uniforms and may not wear certain accessories (e.g., gold necklaces). Another example is regulations regarding absenteeism or late attendance at school. From the teacher's point of view, school regulations are aimed at exercising student discipline and reducing unwanted behaviour, such as reducing absenteeism and late attendance.

We will be punished if we do not attend school without an acceptable reason, come in late to school, and do not wear the appropriate school uniform. (Student, Madrasah A)

We mustn't use any accessories (i.e., gold necklace, bracelets), we have to come to school on time, and have a proper haircut for boys. (Student, Madrasah A).

To make everything happen as expected, for example, the student's absenteeism is getting lower (Headteacher, Madrasah B)

6.2.1.1.3 Safety: stage

The stage dimension in the school climate safety factor refers to the initiation and sustainability of the implementation of school safety regulations. The study found that the four schools had long implemented their safety policies. Then socialisation of the safety regulations often takes place as students start their school. Taking into account the similarities across the four schools studied, it can be concluded that there is no difference between high and low performing schools in this safety factor.

Yes, it (student rules and conduct) has been applied for a long time and consistently applied (Headteacher, Madrasah A).

It has been a long time; the rules have been applied since I started my role as teacher in this school (Headteacher Madrasah B).

6.2.1.1.4 Safety: differentiation

Safety differentiation refers to the extent that safety regulations are applied in the same or different ways to school members. The study found that all four schools consistently applied regulations in the same manner to all school members. All school members are obliged to comply with the safety regulations.

All school members must comply with the enforced safety regulations. The rules are applied to all, and no favouritism (Students, Madrasah A)

We apply the rules equally to all (Headteacher, Madrasah B)

However, the consequences for offenders are flexible. For example, the penalty for students who arrive late is for them to run around the school field. However, this punishment will not be given if the student is in an unhealthy condition.

Apply to all, look at the situation and conditions, but we differ the consequence in some cases. If a student comes to school late, for instance, some students might be asked to run around the school sites, but if they had health problems (i.e. their legs are injured), they would not have been treated the same way, even though some other students were late in [at] the same time (Teacher, School C).

Overall, safety factors of school climate were not different across four illustrative case schools in term of its frequency, focus, stage, and differentiation. This finding is quite similar to the finding in quantitative phase which also found that school safety factors (i.e., teacher safety – see Model 6) had no significant relationship with student achievement. Additionally, student emotional safety (not being bullied) had significant negative effect on student learning outcomes. The finding in this qualitative phase may strengthen that finding (in quantitative phase), because students in all schools, whether in low and high performing schools had similar perception about their school safety.

This finding is inconsistent with most of the empirical research which found school safety as one of school climate factor that influences student learning outcomes (Cohen et al., 2009; Thapa et al., 2013; Wang & Degol, 2016; Wang et al., 2014). This may be due to the fact that much of Indonesian teachers are not aware about the many forms of bullying (Lai et al., 2008; Widayanti & Siswati, 2009).

6.2.1.2 Relationship factor

This section discusses the relationship factor of school climate and its evaluation dimensions. To be specific, the relationship (among school members) outside of the classroom. The relationship factor focusing on relationship/interaction in the classroom (classroom climate) is explained in the next section (differences between high and low performing schools).

6.2.1.2.1 Relationship: frequency

Frequency refers to how often positive interactions take place in school. In general, school members in all four schools similarly indicated that positive interactions did occur. However, it was unclear how frequently/often these positive interactions happened. Most of the teacher-student interactions outside the classroom were informal. This interaction was reflected in some practices. For example, the common practice that students do to their teacher is to greet their teachers when they see their teachers in the school, by saying, for instance, "*assalamu'alaikum"* – (peace be upon you), "good morning" (depending on the time), or kissed their teacher's hand as well as bowing their body in front of teachers (field note from observation). Those common practices (as part of Indonesian culture) are the sign of respect, politeness, and appreciation to their teachers or older people (Rachmadiana, 2004; Saxebøl, 2002).

In addition, other than specific culture, the students also perceived their teachers as kind, nice, respectful, friendly, helpful and willing to help and having fun and socialised with the students. Also, when students want to talk and discuss something, like particular personal problems, most of the teachers were easy to approach.

> *In general, our teachers are kind, respectful and insightful. Sometimes, we do informal activities (i.e., sports) together with them. (Student, Madrasah B)*

They are easy to get in touch with; we enjoyed laughing together with them in an informal circumstance; they are generally kind and caring. (Student, school C)

The teachers also felt that they are well-connected, enjoy talking with their students and treat them as family members. The teacher also maintains a warm and friendly environment in the school, as described below.

The relationship between our student and us is mostly as a family, warm and friendly (Teacher, Madrasah A)

We and our students have good relations and interaction, and we enjoy talking with them (Teacher, Madrasah B)

6.2.1.2.2 Relationship: focus

The focus of student-teacher interaction outside the classroom was mostly similar across the four schools (see quotes below). The focus is to create a pleasant feeling in school, to help students feel good at school. For example, students with family or personal problems can reveal their worries to their teachers (Eccles et al., 1993b). Teachers argued that positive communication between teachers and students plays a key role in successful teaching and learning, the student might do better in their academic task. That teacher view is relevant with some studies (Koplow, 2002; Roeser et al., 2000; Roeser et al., 1996) which found that positive student-teacher interaction increased sense of belonging and thus more academically successful.

…there is a border when a student approaches a teacher, but they are welcome if they need to talk about something personal. The student may choose freely which teacher they want to talk. The informal discussion was about student's personal issues (i.e., family, student-student interaction, etc) as well as academic issues. … (Teacher, Madrasah B)

In term of teacher-teacher relationship, to some respect, the interaction across four schools also had a similar pattern. Most of the teacher perceived that they have a good interpersonal relationship and can communicate with each other easily. They also observed that they have a stable relationship with other school members. However, the relationships were intended more to create a social community and less to develop their expertise in the teaching-learning process and improve student learning outcomes as demonstrated in their practices.

We have trust and can quickly connect with each other. We consider our professional relationship as a family. For example, when one of us is ill; we then visit him/her together. (Teacher, school D)

We do a regular informal meeting every three months to maintain our cohesiveness. If one of us is ill, we will see him/her. (Teacher, Madrasah B)

Some teacher-to-teacher interactions was related to develop their teaching skills as *Madrasah* B did, and try to help the student to improve their academic success as *Madrasah* A did. However, those activities were not well organised, and most were not part of official school planning and policy. All the four schools did it irregularly.

We do help another teacher to teach each other, for example, coaching other teachers on how to deliver a good lesson, but it has only been done occasionally and only done by person-to-person, not part of school policy (Teacher, Madrasah A)

We help each other, a teacher who has good skills in IT, for example, helps others who have fewer skills… (Headteacher, Madrasah B)

6.2.1.2.3 Relationship: Stage

Stage refers to the period at which the factors take place. It is expected that the factors need to take place over a long time to guarantee that the factors have a constant effect on student learning (Creemers & Kyriakides, 2008). In this dimension, all the schools have also reported a similar pattern of the stage, where the relationship factor (good relationship and interaction) in their school has already been going on for a long time.

This good relationship has occurred since we are in grade 7 (Student, Madrasah B)

It has been going on for a long time, since my first duty as a teacher (Teacher, Madrasah A).

And this habit has been going on for a long time (Teacher, school D)

6.2.1.2.4 Relationship: quality

Interpersonal relationship quality is not only about how the relationship benefits from improving student-learning outcomes. It goes further than just being informally aware with others to sharing with and gaining an understanding of one another. A strong personal teacher-student relationship, sufficient personal dialogue, helpful guidance and praise is likely promoting trust. This relationship may lead the student to participate more in learning, behave better and extend higher academic level and also lift eagerness to learn (Rimm-Kaufman & Sandilos, 2011; Roeser et al., 1996). As discussed in the frequency and focus section on the relationship among school members. It was observed that all school have to some degree a good quality of an interpersonal relationship. The good relationship was aimed to help students to achieve more in their academic life.

To me, if in the classroom, my principle is "I'm your teacher", and should be respected as we are. However, when outside the classroom, "I am your friend".

Therefore, if they have a problem, they could approach us quickly to have a heart-to-heart chat personally. Our students also treat us as their parents…if we infrequently make personal conversation with them, it is harder for us to lead them. (Teacher, School C)

6.2.1.2.5 Relationship: differentiation

Differentiation refers to in what way school members connect each other differently in their daily interaction to respond to individual needs. All school have a mostly similar attitude or behaviour to differences in student needs or anything regarding their differences. School treated all members equally.

There is no differentiation in teacher-student interaction, and all members are treated equally. (Teacher, Madrasah B)

We handle all students similarly (Teacher, School C)

In summary, for the relationship factor of school climate can be concluded that there were no differences between high and low performing school. Again, this finding also consistent with phase 1 of this study (quantitative) result, which also found that there was no significant relationship between relationship factor (i.e., teacher-teacher relationship, see Model 6) with student learning outcomes, particularly in the academic achievement of student. Another explanation can be related to Indonesia culture. As explained in Chapter 2 about Indonesian culture, it is argued that the Indonesian culture (i.e., harmony, collectivist minded) may constrain the working relationships of people in educational institutions, including schools (Dardjowidjojo, 2001). According to Hofstede (1986), the society with high power distance index tends to make everything harmonise, for example, students are not expected to challenge their teacher's views. Therefore, maintaining a good relationship is a must in this society, and it may not relate to school learning outcomes.

6.2.1.3 Teaching and learning: frequency

6.2.1.3.1 Classroom climate: frequency

Teachers in all four schools expressed their willingness to motivate their students to learn. Although the teachers pointed out about using a variety of teaching methods, what they meant by a variety of teaching methods is limited to a change of classroom settings. For instance, a class would sometimes be conducted at the school library or laboratory. Although they varied their classroom settings, the qualitative data did not inform any particular changes in their teaching methods.

We applied various teaching methods, for example, we did discuss this week, and for the next week we would carry out something different, going to the school laboratory… (Teacher, Madrasah A)

Do the variation of teaching method (to engage student). (Teacher 6, School D)

Most of the teachers from the four schools used identical approach and strategies in delivering their lessons (i.e., lecture and discussion). As have been noted by past studies, judging based on the frequency evaluative dimension, teaching practices between the high and low performing schools often look similar (Creemers & Kyriakides, 2008). In the next section, the differences between high and low performing schools could be differentiated using the other four evaluative dimensions of DMEE.

6.2.2 Differences between high and low performing schools

6.2.2.1 Safety

6.2.2.1.1 Safety: quality

In this section, safety refers specifically to clear rules and how the school responds to violations of the rules as well as physical safety. The study found that the main differences in

the safety factor of the school climate is in the quality dimension. The quality dimension of school safety associated with activities or policies that have been applied to ensure that school safety is functioning as it is. In *Madrasah* B and School D, the consistency of the compliance system that implemented when students violated the rules is better as opposed to *Madrasah* A and School C. For example, in *Madrasah* B, they have clear rules how to respond to the violation of the rules.

To make sure that student follows the school's rules, the school frequently do a patrol (done by a teacher), searching for a student who breaks the school rules. For example, (one of the rules) on Friday and Saturday, the student must wear a specific uniform, if a student does not do so, she/he will be chastised, and the point is given to the student and is noted on their 'behaviour notebook'.

(Student, Madrasah B)

If a student violates the school rules, In the first place, it might be handled by the teacher or subject teacher, if we are dealing with subject-related problems. If it cannot be solved. Then the student's problem may be undertaken by the responsible teacher. If the problems continue and still have no agreed solutions, the student needs to meet a counselling teacher. (Headteacher, school D)

On the other hand, in *Madrasah* A and School C, most of the mechanism is like a convention with no specific method of addressing the problem. The process and practices are mainly similar to other schools in Indonesia. It indicates that they do not have their own rules for dealing with the offenders.

If a student violates the rules, the first process is by advising them. If the student still makes the same mistakes, then the parents will be invited to the school for discussing the problem and possible solution. Also, the last decision will be

made (removed the student from school) if the student persistently breaks the rules (teacher, Madrasah A).

For example, if a student arrived late at the school, the students have to do some gardening work like pulling out the grass or do some cleaning activities like sweeping the office and or school field, etc. (student, School C)

In term of school safety, *Madrasah* A and School C typically have common characteristics. For example, in School C, because the school is located in a less inhabitant area and do not have a security system, the school was frequently targeted by thieves, and for that reason, some of the valuable school's belongings that support teaching-learning process like LCD projector or computer is stored in the teacher home.

Our school was frequently targeted by thieves because we do not have a dedicated security force. The thieves easily broke the door. We have reported this to the police, but they did not do anything. Therefore, we stored our valuable things in one of the teacher's home. (Teacher, School C)

Madrasah B and School D, on the other hand, have better physical security management. Both schools have security guards that discourage criminals from making school less threatening.

We have two security guards who work until 5 pm and a person who is in school for the entire night. (Teacher, Madrasah B)

Our school are safe and have no theft report, we have security guards. I have also directed a vice headteacher to monitor the safety of schools and students periodically (Headteacher, School D)

This quality dimension of safety can clearly differentiate between high and low performing schools. Therefore, this finding now in line with most of research on safety factor of school climate which mainly found that safety school had significant influence on student learning outcomes (Bradshaw et al., 2014; Cornell et al., 2016; Kutsyuruba et al., 2015; Thapa et al., 2013; Wang & Degol, 2016)

6.2.2.2 Teaching and Learning

In the previous section (similarities between high and low performing schools), it was argued that using only the frequency evaluation dimension of teaching and learning could not differentiate between high and low performing schools clearly. In this section, the result found that the key differences between high and low performing schools are in the other evaluation dimensions (Creemers &Kyriakides, 2008), particularly in focus and quality.

6.2.2.2.1 Teaching and learning: focus

In data analysis, it was found that focus on teaching-learning can be categorised into two aspects: academic and non-academic. However, the most differences between high and low performing schools were on academic focus. Academic is about student performance in their subjects, and on the other hand, non-academic is about student self-development. In School D and *Madrasah* B (higher performing schools), it was found that they have a better focus on academic achievement. High performing schools had a high focus on their student academic achievement. This can be seen from the reason of students when of choosing the schools. Most of them chose the school because of its academic excellence at least in their city and can lead them to get accepted into a good school in their further education.

> *(We chose this school) Because I want to go to Insan Cendekia (one of the very best Islamic high school in the country. Its graduates mostly without difficulty go into any universities). (Student, Madrasah B)*

I chose this school because this school is one of the best schools in the city, have a good discipline orientation. A lot of graduates also go to a prestigious high school (Student, School D)

We are one of the best schools in this city and our school was chosen to run the new curriculum as a pilot school. This is because of our academic excellence (Headteacher, School D)

The head teachers in two schools, on the other hand, often emphasis well on the academic achievement of students. Both head teachers encourage their teachers regularly to track their progress.

…for example, I asked the teacher to provide their students' score on the half-term exam. Then, at the end of the term, the score must be higher... (Headteacher, Madrasah B)

I order all four of my vice headteacher to help the teacher in collecting information of student academic progress and reported back to me once a week, so I know the progress of the students (Headteacher, School D)

Moreover, *Madrasah* B had a program called 'learning clinic' to help their student enhancing academic abilities. The students have opportunities to consult any lesson (particularly maths and sciences) that they may not fully understand when they have been taught in the classroom. To make this program running better, before starting the program, the survey was conducted to analyse the tendency of students to choose their preferred teachers. This program, according to the headteacher, mostly gave a significant impact on student academic learning. The headteacher regularly checked the progress of 'learning clinic' impact in the school meeting. She asked the teacher to pay more attention to students who have grades below the minimum standard.

We have a program called 'learning clinic'. Each student was allocated to a teacher, but beforehand, a survey had been conducted asking students' preferred teacher. Then the student makes an appointment with the teachers to set their learning goal. (Headteacher, Madrasah B)

Relatively identical to *Madrasah* B, School D also had a program to track student progress. The headteacher assigned all teachers to track a group of students to participate in the program. At least 14 students should be monitored by one teacher. The monitoring progress including student academic achievement, personal development and behaviour. Then, teachers must report the progress to the deputy headteacher on a regular basis, and if any problems arise, they will be addressed at the meeting.

So, how students behave, their academic achievement progress will be reported to the vice headteachers weekly. We have four vice headteachers and about 56 teachers, and each vice headteachers have a responsibility for monitoring teacher's activities. So, each vice headteachers was allocated to monitoring about 14 teachers… Next, the vice headteachers will report to headteacher … then, in a monthly meeting, we will discuss student progress. (Headteacher, School D)

Conversely, in a low-performing school like *Madrasah* A, academic achievement was found somewhat less focus on academic performance compared to *Madrasah* B and School D. There was no competitive climate for academic achievement, which is reflected in the low score for most of their examination results. Former deputy headteacher blamed the raw input of the student. He argued that only low ability student enrolled in the school. Also, the minimum number of students was blamed as caused by this unwanted circumstance.

We set-up 75 as a minimum score to pass the exam, but most of our students' exam result was under the preferred standard. They did remedial several times to get the intended result…They do not have a sense of competition. That made us repent…Next, the small number of them make them did not want to do more because they argued that if they became the worst in the classroom, they were still among top five best student because they were only 4 students in 9th grade…

(Headteacher, Madrasah A)

Another aspect that made up this unwanted situation was because of the school regularly marked-up student's final score, which used to determine if a student can be accepted in the next grade.

We marked-up their low score to pass a minimal criterion as a requirement of completion of the study. The minimum standard is a combination of a national exam and school exam. So, their score in school exam was marked-up…

(Headteacher, Madrasah A).

Other lower performing school only had a focus on enabling their students to go into public schools. This public schools based their acceptance on students' national examination score.

They could compete with other students from other schools to get into public schools. At least half of our students went into public schools. (Teacher, School C)

This was happening because before they did the national examination, we (teachers) had given them additional lessons. (Teacher, School C)

Marking up students' grades is unethical. However, there is a logical reason for the teachers to engage in such action. Public secondary school have more qualified teachers, better facilities, and are well financed (Hendajany, 2016; Newhouse & Beegle, 2006). In order to get into such a school, students must achieve a certain pass grade in the national exam (Faisal & Martin, 2019; Furaidah et al., 2015; Indonesia, 2003). School's lack of ability to support their students to fairly attain the passing grade may have motivated the teachers to cheat to increase their students' chance to be admitted in public schools.

Moreover, in a low performing school like *Madrasah* A, they had no special activities to enhance their student academic outcomes. Even the school had no student consultation services which most schools has.

We don't – [shaking their head together] (have counselling teacher) (Student, Madrasah A)

Similar to *Madrasah* A, School C also had no specific activities to support student learning due to limited resources, particularly teacher's workforce. As comparable as *Madrasah* A, for example, the school had no individual teacher acting as a counselling teacher. This function is shared among available teachers. The absent of dedicated counselling teacher means that the ordinary teachers have to spend more time in dealing with student personal problem that affects learning.

Honestly, this circumstance indeed makes us exhausted. We don't have dedicated counselling teacher, so we are also acting as counsellor, even the headteacher. (Teacher, School C)

6.2.2.2.2 Teaching and learning: differentiation

High performing schools treated their students differently based on their abilities. More able students are grouped in specialised classes and are taught by specialised teachers. This

may appear discriminatory, but such specialised classes are meant to enable the more able students to learn faster without having to wait for their less able peer. On the other hand, this also helps the regular class teacher to slow down when teaching the average or less able students. In contrast, low performing schools do not have the means to create such specialised accelerated classes. As such, students in low performing schools, whether they are more or less able, are treated in the same way.

…standard service was applied in a regular classroom, though we had special treatment for a particular classroom. They got special coaching… (Teacher, Madrasah B)

Contrary, *Madrasah* A and School C did not have many choices to do a variation. Therefore, these schools applied the same way to teach students.

No differentiation, they are treated equally… (Teacher, School C)

We did not do any separation; all was taught similarly… (Headteacher, Madrasah A)

6.2.2.3 Classroom climate

Teaching and learning practices mostly happen in the classroom. Therefore, this factor describes the classroom activities of teaching and learning which are associated with classroom climate. As argued in Chapter 3, the classroom climate is part of a school climate that occurred inside the classroom. Creemers and Kyriakides (2008) argued that in general, classroom climate can be divided into two main aspects: (1) classroom interaction and (2) classroom disturbance. Therefore, those two aspects of classroom climate will be described below.

6.2.2.3.1 Classroom interaction: quality

The quality of the teaching-learning process between high and low performing schools was apparently unequal. In the high performing schools, the student said that the way of the teacher conducting the lessons was quite effective. If the student did not understand what was learned, the teacher sometimes checked the student's understanding by walking around the class and helping students. By doing this, student perceived that they could understand what the teacher said, however, students also felt that the time was short.

It was effective (the way teacher walking around to ensure student's comprehension, but the time was too tight and somehow could hamper our accomplishments (student, Madrasah B)

Went to their desk to encourage them, if needed, and asked what had not been understood… (Teacher, School D)

This statement also may give an insight that the teacher was perceived as a helpful teacher. Contrary, the teacher in a less performing school was seen as not helpful to ensure the student received the lesson as expected.

It's hardly ever occurred (because the teacher went to the student's table to explain more when the student did not understand what the explanation was). (Student, Madrasah A)

Though on another low performing school, students could understand what their teacher was conveyed, but when the teacher gave the assignment to be solved, they did not know how to solve it.

We understood the explanation of the teacher and the illustration they gave, but when the teacher gave us the assignments, we didn't know how to answer (Student, School C).

6.2.2.3.2 Classroom disorder: focus

Focus is measured by looking at the specific problem, whether an incidental or a continuous challenge and how good the teacher deal with those problems. It is all about teachers' creativity. Therefore, some teacher, particularly in the high performing school, could manage the classroom efficiently. For example, in *Madrasah* B, the headteacher used to observe how the teacher conducted the lessons. She found that the teacher often did not need to use various teaching activities to treat classroom disorder.

I occasionally observe how the teacher conducted the lessons. Sometimes, the teacher did not have to use various teaching methods to deal with classroom disorder. It entirely depended on teachers (Headteacher, Madrasah B).

At School C for example, the way they deal with problem rise in the classroom is based on their previous experience with the student. However, typically, teachers in *Madrasah* B and school D have various strategies to manage their class. Besides, the teacher in School D and *Madrasah* B created classroom rules by involving students. Teacher in school D also has another strategy to deal with disturbing student, for example, by arranging a student desk. She put the boy and the girl beside one another.

For example, to minimise the bothering pupil, we organise their desk, place the boy and the girl one after the other. (Teacher, School D).

6.2.2.3.3 Classroom disorder: quality

Quality is seen in connection to the impact that the teacher's behaviour has on solving the problems that arise, as measured through students' behaviour. As mentioned before that teacher in *Madrasah* B and school D have more effort to anticipate issues that possibly come up in the classroom by creating classroom rules and scheduled some specific lessons (i.e., math) in the morning. The teacher does this to make student more focus on their lesson, as habituation of discipline behaviour as well as to make student respect each other. On the other hand, in a low performing school, the classroom is less organised. As described in the previous focus dimension, the lack of class organisation at low performing schools may be due to the teachers' lack of skills in managing their classroom.

Students and teachers make the rules of the class together (Student, Madrasah B).

The rules were agreed before the start of the first lesson at the beginning of the new school year (Teacher, School D)

6.2.2.4 Institutional Environment

6.2.2.4.1 School facilities: quality

In general, all school members across all schools perceived that to some degree they have at least minimal school facilities (classroom, laboratory, and teaching materials). However, the facilities need to be improved or were not in good conditions. This is the indication that the quality of facilities particularly in School C and *Madrasah* A. For example, in *Madrasah* A, teachers and headteacher observed that their school have enough resources, but their student has seen that the quality of those facilities was poor. This is also happening in School C, where the teacher perceived that school had enough resources, but the headteacher said that they did not have enough funds.

We have blackboards in each classroom, laboratories, computer, but only one computer, and the blackboard needs to be replaced because it is broken (Teacher, School C).

We do have a library, but when heavy rainfalls down, our school will be full of water and became flooded. This also affected our library. (Headteacher, Madrasah B)

The school environment in these two high performing schools is reasonably clean and well-ordered. School D has won a prestigious award in School Environmental Management from the Ministry of Environment and Forestry of the Republic of Indonesia in term of their concern on creating green school environment. In *Madrasah* B, the headmaster confirmed that that people who visited their school found the madrasah were clean and organised.

Almost all people who visited the school said that our school were clean. (Headteacher, Madrasah B)

The researcher also noted that School C had limited and poor-quality facilities. For example, indeed the school has a library, but its use is limited because the library is damaged, the ceiling is also broken, the roof is leaking, the books are not properly maintained, there is no librarian on duty, the books are not properly organised and dusty, and there are also broken tables. In practice, it can be argued that they did not have a library that can support learning. The circumstance of some classrooms was also relatively similar to the library. They did not have a dedicated person to clean the school. Therefore, cleaning duties were the responsibility of all school members, including student, teacher, staff, and headteacher.

This condition is relatively similar in *Madrasah* A. Student reported that their school's facilities were inadequate. For example, their blackboards were a painted wall, the classrooms were mostly dirty, and classroom's ceilings were impaired. The classroom's wall is almost full

of scribbles. The researcher also witnessed similar problems as expressed by students when visiting the school.

Next, the temperature of the classroom was rather hot, and the light, both natural and electric, was not enough to illuminate the entire class. In addition, both schools reported odorous, dirty and uncomfortable toilets and insufficient supplies of water.

6.2.2.4.2 School engagement: quality

School engagement refers to positive identification with the school as well as broad participation in school activities. Generally, school members across all schools perceived that they like being in school. However, in term of positive identification with school, high and low performing schools have significant differences. For example, teachers in *Madrasah* A and School C reported that they are proud to be a teacher in their school. However, they were not proud of the school's physical conditions.

Actually, we are proud as a teacher. However, we are not really happy with our school condition. We will feel prouder if only this school is moving forward.

(Teacher, School C)

If I am not feeling proud, I will leave the school. I feel honoured in this school.

(Teacher, Madrasah A)

The students have mixed perception about the school condition. Some students like being in the school while others for some degree not really liking their school. The students do like the teacher but not surely proud of being part of the school.

Unlike, *Madrasah* A and School C, school's members of *Madrasah* B and school D perceived that they like being part of the school because the school is among the best in the city. Teachers also said that parents claim that their school is one of the best schools and have various extracurricular activities. Next, particularly in *Madrasah* B, they were really proud and

felt pleased because the school is not only teaching academic matters but also teaching Islamic values.

We are the oldest and the most favourite madrasah in the city. (Headteacher, Madrasah B)

We have a lot of activities, both academic and non-academic. We have 23 extracurricular activities and have an agreement with other institution to involve our students in various activities. (Teacher, Madrasah B)

Our school ranked consistently in 3rd or 4th place in Padang city. In the province level, we are on 6th place out of about 1300 secondary schools. (Headteacher, School D)

Also, the relationship between school and their graduates is strong. To maintain this strong emotional connection, some activities have been made. For example, graduates from *Madrasah* B would come back to their school to see their former teacher and ask for their blessing when they about to take the national exam in the senior secondary school. Also, when the researcher came to school, there was an activity that aimed to support orphan students in the school. They pray together and then give charity to the orphan student. This is an annual ceremony of the school as the expression of Islamic practice as taught by Prophet Mohammad to take care of orphans (field note).

6.2.2.5 School improvement process

6.2.2.5.1 School improvement process: quality

School climate improvement efforts aim to enhance individual learning and behaviour (Thapa et al., 2013). Bryk et al. (2010) argued that there are four main areas of concern for school improvement: (1) professional capacity development focusing on teachers' skill and

knowledge, (2) improving orderly, safety, and norms of the school, (3) involving parent, and (4) focusing on curriculum alignment.

In the high performing school, the headteacher made a good effort to improve their school performance. For example, in *Madrasah* B, the headteacher gives the direction and monitoring the student progress frequently.

> *We have a regular meeting to monitor student progress. In that meeting, I ask the teacher to familiarise with their students well, so they know what progress they have made so far. (Headteacher, Madrasah B)*

To monitor order, safety, and the headteacher usually walking around the school and if she found something interesting, she will take a digital picture using her smartphone and share it in a school meeting to overcome the problems.

> *I take some pictures to prove some unwanted situation, including the teacher behaviour in the classroom and then share the image in the meeting to solve the problems (Headteacher, Madrasah B)*

Compare to a lower performing school, for example, in School A, the headteacher seems to not have the power to manage the school. She also does not have leadership skill to manage the school better. When she wants to evaluate the teacher, she only uses the standard form from the local government, almost no creativity.

> *There are points from the education office, I used this to evaluate the teacher (Headteacher, School C)*

6.2.3 Differences between *madrasah* and general school

6.2.3.1 Teaching and learning: Focus

The main differences between *madrasah* and general school are mainly on the non-academic focus of schooling. *Madrasah* had greater focus on practising Islamic values, particularly the practices that most of Muslim cannot do. For example, the school aimed to prepare the student with the skill of Islamic funeral services, which is considered as community obligatory in Muslim society. This practice is including washing, shrouding, do praying for, and burying the body. In Islam, it is highly recommended that their sons or relatives should do those activities and be the leader (*imam*) of funeral praying.

We encourage our student to learn one of important community obligatory in Islamic practices. At least they can pray for their parents as well as pray for them when their parents passed away. (Headteacher, Madrasah A)

This school has better religious values than general schools (Teacher, Madrasah B)

Contrarily, in general schools, regarding non-academic focus, the school did not have unique treatments. They only did formal activities as written national curriculum. In the curriculum, the non-academic focus is on character education. For example, before the lesson started, student and teacher pray together or encourage the students to be confident to give speech in front of their friends.

We do teach about character education by praying before starting the lesson as an expression of gratefulness we receive from God. This is as printed in our lesson plans (Teacher, School C)

For example, when it's time for short speech after praying, students alternately give a speech to other students… (Headteacher, School D)

The case study schools of high and low performing schools also can be differentiated to its status. In this study, the high performing schools (*Madrasah* B and School D) are public school, and the low performing schools (*Madrasah* A and School C) are private school. Therefore, the main differences are due to its status. It has been argued before that public secondary schools have better resources, including teachers, building, learning material, supporting facilities like science laboratory and language laboratory. Therefore, many studies in the Indonesian context found that public school is outperformed by private school (Hendajany, 2016; Muttaqin et al., 2019; Newhouse & Beegle, 2006). Indeed, some other private secondary school is better than the public one as in Bedi and Garg (2000) research report. However, the better private schools are mostly an expensive school and only student from high economic family background could be accepted.

Other research also found that there are no significant differences between *madrasah* and general school (Muttaqin et al., 2019; Newhouse & Beegle, 2006) and the differences were mostly determined by school resources including highly qualified teachers, funding as well as school physical resources (Ali et al., 2011; Muttaqin et al., 2019). Therefore, this research result, particularly in this qualitative phase, is not in agreement with the research by Parker and Raihani (2009) which concluded that *madrasah* student tends to have lower achievement compared to student in general schools.

6.3 New potential factors that emerged from the data analysis relevant to highlight differences between the high and low performed schools

RQ2.2 focusses on new factors that can be obtained from a qualitative inquiry that are relevant to highlight differences between the high and low performing schools. There are two keys relevant factors that emerged from the data analysis as described below.

6.3.1 The orientation of education

This theme emerged from the qualitative data analysis process. The education orientation of the four schools was mainly academic which can be concluded from their emphasis on the examination orientation. Orientation of education means to what extent do stakeholders prioritise testing within the education system. This view might lead all related stakeholders including teachers, headteachers, and policymakers in delivering education system, what is seen as an essential aspect of education, how to achieve it and how to do an evaluation of learning outcomes. This study shows that despite the Indonesian education system has been reformed many times, but in reality, the learning outcomes are classically measured by how good student doing in the exam. Besides, school always put their student in rank based on academic performance. Therefore, it is not hyperbolic if it is believed that the orientation of education in Indonesia is mainly to train students to pass tests and get good grades (Effendi & Suyudi, 2016). This orientation also reinforced by school selection system onto the next education level that overlooks on students' score in the national exam. Likewise, the parents, who want their children to go to favourite or best schools, give more lessons by sending their children to after school academic tutoring agency or invited ones who can teach additional lessons to their children at home (Thahir & Hidriyanti, 2014). This education orientation named by Armstrong (2006) as an academic achievement discourse. This practice can be clearly realised in one policymaker comment below.

To do an intervention in the school in the next following year, in what subjects were student got the lowest score, for example, maths, therefore (the intervention) will be focused on teacher training in maths (local policy maker)

Teachers also felt that test-oriented education led them to focus only on delivering curriculum contents without having much opportunity to develop student competency in a subject they carried out.

We were chased by the local government to participate in the council-led exam (together with other schools in the city for half and end semester term). In fact, not all curriculum content had been delivered yet. Therefore, we did need to prepare our students with all intended curriculum contents. Because of this condition, most of the students did not really understand what they got. (Teachers, school D)

The policymakers in the Ministry of Education have realised negative sides of testing orientation education. Most schools only teach the subjects tested in the national exam to students. This led students to conclude that the main purpose of education is only to pass an exam and achieve high grades. A heavy reliance on grading and testing to determine whether a student has been successful or unsuccessful in learning, according to Armstrong (2006), is a strong indication that the education system is not concerned with social-emotional growth of students.

Therefore, in general, schools only teach students the subjects assessed in the national exam. It led students to think that the goal of education is only to pass an exam and get good grades. (Policymaker at MOEC)

Exam-oriented education is not only happening in the Indonesian context. Other countries like UK, US, China, Tanzania, and Kenya also have problems with this exam-oriented education (Jennings & Bearak, 2014; Kirkpatrick & Zang, 2011; Mackatiani, 2017; Marshall, 2017; Salim, 2011). Exam-oriented education influence teacher's instructional practices and methods (Jennings & Bearak, 2014; Mackatiani, 2017), therefore, it influences the classroom and school climate as a whole. It is good to emphasis on academic attainment as many school climate researchers argued (Cohen et al., 2009; Loukas, 2007; Voight et al., 2013; Wang & Degol, 2016; Wang et al., 2014; Zullig et al., 2011), but teaching practices that only

oriented to teach student passing the test is the worst practices (Furaidah et al., 2015; Kirkpatrick & Zang, 2011), since it can negatively influence critical thinking (Mackatiani, 2017; Revina, 2017), as one of main aim of education.

6.3.2 School culture

The other theme that emerged in the data analysis process is school culture. School culture is one of an essential aspect of school climate in Tagiuri's model (Tagiuri, 1968) and the model supported by Anderson (1982) who published the first review in school climate research. Tagiuri's model proposed a taxonomy of school climate including ecology, milieu, social system, and culture (see Chapter 2 for detail).

In this research, the differences between *Madrasah*s and general schools are mainly related to school culture. It is unwritten rules and traditions, norms and expectations and the group's shared behaviours. Moreover, Peterson & Deal (2009) stated, "Culture consists of the stable, underlying social meanings that shape beliefs and behaviour over time" (p.7). Indeed, the culture differences because those two types of schools applied different bases as well as supervised by the different ministry. *Madrasah*s base on Islamic values, whereas general schools base on general Indonesian values.

In public schools, the student learns only two hours of Islamic Religious lesson per week. Contrary, in *madrasah*s, student need to learn about five to six hours per week plus all subjects as taught in general schools. The student in *madrasah* has additional lesson such as *Quran* and *Hadith studies*, *Fiqh* (jurisprudence), *Aqidah* (theology), *Akhlak* (virtue) as well as Islamic History (Tan, 2014b).

General schools have nine subjects, whereas madrasahs have fourteen subjects. It includes all lessons as taught in general school and specifics lessons including aqidah ahlaq (theology and virtues), Quran and hadith studies, fiqh (Islamic jurisprudence), Islamic History, and Arabic language. 2 hours for each

subject. So, they learn much more than their peers in general school.

(Policymaker at MORA)

Also, students' motivation when choosing *madrasah*s supports the culture difference between the two types of schools. Of students' or reasons chosen *madrasah* is to get more in Islamic studies and practices, discipline in prayer, as well as to make a deeper understanding of Islam.

6.4 Chapter summary

This chapter reports on the findings of analysing the qualitative interview and focus group data from four Indonesian case study schools and four policymakers. An analytic framework on DMEE and Thapa et al (2013) was used to create a priory code to analyse the data. By using this combined framework, it was found that this approach was useful in analysing the qualitative data. There are no previous qualitative studies that conducted using this effectiveness dimension, most of the studies were conducted using quantitative methods (i.e., Kyriakides & Creemers, 2008a; Kyriakides & Creemers, 2009; Creemers & Kyriakides, 2010; Panayiotou et al., 2014).

However, by utilising the evaluation dimensions of DMEE to explore and analyse school climate, robust differences between school climate of a less and high/low performing schools can be found as well as different types of school as in the Indonesian context (general school and *madrasah*). In this research, by only measure school climate using ordinary factors, it cannot differentiate clearly how school climate influences school effectiveness. The differences between less and effective school are mainly on the focus and quality dimension of DMEE.

However, the effectiveness dimension of DMEE cannot fully apply to all school climate factors. For example, in the emotional safety factor, it is quite complicated to use all effectiveness evaluative dimensions since the factor asking about how school members

perceived their emotional safety in the school. Though, this research attempt to measure school climate differently to give more idea on which part of the school climate should be improved in a journey to shape an effective school.

This study also found other factors that might better explain the assessment of the school climate in the Indonesian context or in other countries that may have a similar educational system or culture. The factors that emerged are education orientation and school culture. The first is a new factor that never been considered as school climate factor in the existing literature and the second gives emphasis that school culture is part of school climate (see Anderson, 1982; Tagiuri, 1968).

Chapter 7: Discussion and Conclusions

7.1 Introduction to the chapter

The present thesis examined the impact of school climate on students' math learning outcomes. Two general research questions have been analysed using quantitative and qualitative inquiry. This chapter will end this dissertation by outlining and discussing the main results and implications to reach a general conclusion, and therefore to address Research Objective five and six. Finally, the chapter will identify the main strengths and limits of the study and will end by underlining areas of further research.

7.2 Discussion and implications of key findings

7.2.1 Discussion of key findings of RQl: What are the differences of school and classroom performance in Indonesian lower secondary schools in terms of mathematics and self-beliefs? If such differences exist, to what extent does school climate predict the differences?

7.2.1.1 Key finding 1: The range and extent of students' math performance and self-beliefs among Indonesian Year 8 students

The null model results show that around 34% to 36% of the total variance in MATH, KNOW, and APPLY are attributable to differences between schools. Interestingly, REASON, which is also a part of math cognitive domain had less variation attributable to differences between schools. Only about 26% of REASON score variance is attributed to school differences. The REASON element of TIMSS's math achievement is related to the capacity of logical thinking, ability to conduct intuitive and inductive perceptive to reach solutions to non-routine problems (Mullis et al., 2009).

Lower REASON variance at school level may indicate that schools in Indonesia are the lack of promoting higher-order thinking. The teacher spent less time on nonroutine problem-

solving in math classroom (Ragatz et al., 2015). Ragatz and colleagues (2015) argued that the condition may be affected by the National Examination orientation which much more endorsed the use of memorisation and routine problem-solving.

In term of classroom differences, the variation of math and its cognitive domains between class within school was relatively small. The variation was only around 8% to 9% MATH, KNOW, and APPLY variance are attributable to classroom differences. This small classroom differences even lower in REASON. It is approximately 5% can be attributed due to classroom differences. The low classroom differences in explaining learning outcome as found in this research are contrary with other studies which suggest that difference between classrooms compare to variation between schools is much more substantial (Hill & Rowe, 1996; Rowe & Hill, 1998).

However, that result is not fixed and sometimes inconsistent and open to a wide range of interpretations (Hill & Rowe, 1996). In this research, the low variance between classroom might be because of its low variation of classrooms sample in the TIMSS study. In the Indonesian context, as mention in Chapter 4, it is only 20 schools that have two classrooms, the other 54 schools only have one class participated in the study.

Unlike math and its cognitive domains, in the null model, self-beliefs have a smaller variance that can be attributed to schools. For SELFCONCEPT, the differences between school are only 5%. On the other hand, for SELFEFF, there are slightly better significant differences, about 10% is attributed to school differences. However, the differences between class within the school are also small on self-beliefs outcomes compare to math and its cognitive domains. Classroom variance on this self-beliefs' outcome was about 4% for SELFCONCEPT and 6% for SELFEFF. These finding resonated other researchers that mainly found that the differences between schools or classrooms were more noticeable on academic achievement rather than in self-beliefs or affective outcomes, for example, research conducted in Belgium (Opdenakker

& Van Damme, 2000), UK (Gray, 2004; Thomas, 2001), and Cyprus (Creemers & Kyriakides, 2010).

The proportion of variance attributable at the school level reflects the achievement gap in student's math performance in Indonesia, which is similar to other developing countries (e.g. Zanzibar (Yu & Thomas, 2008)). However, the average school-level variation in academic self-beliefs was low. Overall, Model Null showed that 35% math performance in TIMSS 2011 results can be attributable to the effect of schools. Thus, a significant achievement gap can be described in the Indonesian school system. This result is somewhat higher than previous estimates of primary school effectiveness research in Indonesia two decades ago (Kaluge, 1998). Kaluge (1998) found that 29.2 per cent of mathematics in school variance was attributed to the school.

However, the pattern of raw results was relatively lower compared to other developing countries such as Brazil, Colombia, Honduras, Egypt, India, Jordan, Namibia, Pakistan, Thailand, Zimbabwe, Botswana and Philippines, where the average school-level variation has been reported as 46% at the primary level, and 41 % at the secondary level (Riddell, 1997). In another context, Yu and Thomas (2008) reported that Zanzibar had relatively similar achievement gap with 34% variance in math performance was attributed to the school. Contrary, in the developed country such as the UK, has shown that only 14% of the total unadjusted variance could be attributable to schools (Thomas & Mortimore, 1996).

7.2.1.2 Key finding 2: The range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs after controlling for student characteristics.

Considering student background variables 34% to 35% of the variance in MATH, KNOW, and APPLY is attributable to the difference between schools, and 26% in REASON. Moreover, differences between class within the school were respectively similar. For self-

beliefs outcomes, this proportion of school differences in SELFCONCEPT and SELFEFF also similar to variance component null model. This result indicates that, in the Indonesian context, student background variables did not have a significant effect on explaining the observed differences between schools, classrooms or students in student outcome learning. However, it is must be noted that the finding did not include student previous achievement, therefore should interpret this finding carefully. In many SER study, the previous attainment is the most influential student-level variables (Gray et al., 2001; Lenkeit, 2013; Muijs & Reynolds, 2003; Thomas, 1998; Timmermans & Thomas, 2014).

This study found that gender only has a statistically significant relationship with KNOW and SELFEFF, but in a different direction. For KNOW relationship with gender is negative, where boys have typically lower KNOW score compared to girls. Conversely, the boys have a higher SELFEFF score than girls. The gender effect on student learning outcomes in previous studies was also inconclusive. Some studies have found that there are gender differences in learning outcomes (Muñoz-Chereau, 2019; Sammons et al., 1993; Thomas, 2001), while others found that the gender did not have influenced on learning outcomes (Strand, 2010, 2016).

On other student's socio-cultural factors, parent education and the number of books also found to have a significant relationship with student achievement and self-belief. In term of parent education background, generally can be concluded the lower the education of the parents, the lower achievement of the student. However, parent education background had no relationship with student' self-beliefs. In previous research, the parent education had a positive influence on the students' academic achievement. Students of parents with higher levels of education performed better in academic than students with lower levels of education (Gooding, 2001). Acharya and Joshi (2009) in India, also reached the same conclusion after they examined the relationship between academic achievement of two hundred teenagers (16 to 19

years old) and parental degree of education. The possible assumption is that parents learn something at school which affects how they communicate with their children through home learning and the capacity for children to model (Eccles, 2005).

In term of self-beliefs, this study found that parent education background had no relationship with student' self-beliefs. The potential explanation can be found in studies by Eccles (2005), which proposed that higher education parents may have strong or disproportionate parental control that can contribute to decreased student confidence.

For study support, compared to a student that has better study support (have room and other goods in-home), the student that have less study support perform lower in MATH, KNOW, and APPLY, but not in REASON. However, the relationship between home study support and SELLEFF was opposite as in math achievement. The student who has either own room or other goods had positive SELFEFF, but for SELFCONCEPT, there was no relationship with home study support. An empirical analysis by Filmer and Pritchett (1999) concerning home study support also found that there were various trends between student wealth disparities and academic achievement across countries. The relationship between household resources was therefore inconsistent in this study as well.

In term of the number of books at home, there is an inconsistent relationship between the outcomes. For example, having fewer books in comparison to no books had a negative association with MATH and KNOW. On the other hand, there was no relationship with APPLY, but having between 26-100 books had a positive correlation with REASON. However, the trend was mostly positive with this cognitive domain, indicating that the more book they have, the higher REASON score they have. This result is also similar to SELFCONCEPT and SELFEFF, which found that the number of books has a positive relationship with the student's self-belief (see Table 5-2 for score detail).

Moving to another socio-cultural background variable. The study uses language (how often students use the language of the TIMSS test (Indonesian language) in their daily life) as a proxy to measure ethnic differences. The result found that student who never uses the language had a significant negative relationship with REASON, but not on other achievements. The trend was also similar to self-beliefs outcomes which the study found that also have a negative relationship with language use. The previous research on ethnicity and learning outcomes also inconclusive (Strand, 2016; Worrell, 2007).

The study showed that SELFCONCEPT has a substantial positive relationship with MATH and its entire cognitive domain in terms of the psychological context variables of the student. Math achievement, on the other hand, also had a strong positive correlation with student self-belief. This result is in line with most of the research in the field (Marsh, 1990a; Marsh & Martin, 2011; O'Mara et al., 2006; Pajares & Urdan, 2002; Parker et al., 2014). This study therefore confirms the reciprocal relation between achievement and self-beliefs (Huang, 2011; Marsh & O'Mara, 2008; Seaton et al., 2014).

The percentage of total variance explained by student background variables were relatively low, only 5% to 6% of the total variance in MATH, KNOW, and APPLY (see Table 5-2). For REASON and SELFCONCEPT, the total variance explained was 4% and only 1% for SELFEFF. This result suggests that considering student's characteristics variables in the model was statistically significant but nevertheless hardly adequate to predict students' learning outcomes, because the percentage was low.

7.2.1.3 Key finding 3: The range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs after controlling for teacher's characteristics.

The relationship between math achievement and all teacher's characteristics was statistically significant, except for gender. In term of the self-belief outcome, however, none of

the teacher variables had a statistically significant relationship with student's self-beliefs. In previous studies, the relationship between teacher's gender and student achievement were inconsistent. For example, teacher gender found as a non-significant teacher variable in Antecol et al. (2015) study who examined the effect of teacher gender on primary student's achievement in the US using a randomised experiment. In developing country context, like Pakistan, teacher gender had a strong relationship with the student's performance (Warwick & Jatoi, 1994).

Teacher age had a significant negative relationship with MATH, KNOW, APPLY, and REASON. The older the age, the lower student's achievement. This result is also similar to the teaching experiences variable. The teacher who had lesser experience tend to have lower student achievement. This result is consistent with most of the research in the area which also found that teacher experience and qualification had a significant positive effect (Croninger et al., 2007; Darling-Hammond, 2000a, 2000b). This situation also applies to the teacher's educational background. The student who was taught by a teacher who had no formal education beyond upper-secondary tend to have lower math score and its cognitive domains. The trend was the higher of teacher education and the more specialist in math, the higher student achievement.

7.2.1.4 Key finding 4: The range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs after controlling for school characteristics.

After controlling school characteristic which include school's social and economic background (aggregated from student's SES), school location, school size, and type of school (general private school, general public school, private *madrasah*, and public *madrasah*), this study found that as expected school with higher average SES tend to have significantly higher achievement in MATH, KNOW, APPLY, and REASON. This result also reinforced previous

studies that have stressed the importance of school's SES on academic achievement (De Fraine et al., 2002; Opdenakker & Van Damme, 2005; Opdenakker et al., 2002; Sammons et al., 1994; Timmermans & Thomas, 2014)

Regarding school location, students who attend schools in the small town achieved significantly lower than their counterparts in remote rural, medium city, and suburban, and urban. This result also in line with previous research such as Young (1998) and Burger (2011), who also found the achievement differences between urban and rural students. However, the difference is quite unusual, because the student in remote rural had higher achievement compared to a student in small town. This finding is also somewhat comparable with research conducted by Tayyaba (2012) in Pakistan, which found that rural and urban students were equally successful in academic achievement across some provinces. In Malaysia (Indonesian neighbour country), Othman and Muijs (2013), found that there was no gap in school performance between rural and urban school.

Moving on to school types, also as expected, the student who attended private schools, both *madrasah* and the general school had significantly lower achievement. This finding also in line with previous research conducted by Hendajany (2016). She utilised the Indonesian Family Life Survey (IFLS) data, which found that public school student had a higher achievement level compared to student who attends private schools. Comparing *madrasah* and general school, this study also supports that there is no achievement no differences between private *madrasah* and private general school as study conducted by Newhouse and Beegle (2006) in Indonesia which also used IFLS data.

Unlike the cognitive achievement outcomes, SELFEFF had a significant negative relationship with the school's average SES but had no relationship with SELFCONCEPT. For school location, a student in suburban had a significant positive relationship with their self-beliefs. Concerning school types, self-beliefs had no significant relationship SELFCONCEPT

and SELFEFF. There is very limited previous research that looked at self-beliefs relationship with the school context.

The variability of math score and its cognitive domain between school has considerably dropped up to 16% compared to Model 0 in MATH, KNOW, and APPLY. For REASON, the variability between school also dropped from 26% to 15%. In term of classroom variability, the trend remained similar to Model 0 with slightly increased up to 2%. Consistent with previous research which mostly found that when the school context is taken into account, the apparent school effect is reduced (Muijs & Reynolds, 2003). However, unlike achievement outcomes, SELFCONCEPT and SELFEFF the variability between school and classroom was almost identical with the variance component model, with 2% in SELFEFF.

By taking into account the school context, the goodness of fit of the model was significantly improved, explaining 16% to 17% of the total of variance in MATH, KNOW, and APPLY. For REASON, the total variance explained was slightly lower (12%) compared to other cognitive domains. However, for SELFCONCEPT and SELFEFF, the change was minimal. This goodness of fit is considerably better to predict students' learning outcomes. It suggests that school variables are better than student and teacher background variables in explaining student achievement.

7.2.1.5 Key Finding 5: The range and extent of school and classroom performance among Indonesian Year 8 students in math and self-beliefs after including school climate factors and controlling for student, teacher, and school characteristics.

Taking into account all significant school and classroom climate factors and controlling for all significant student, teacher, and school characteristics, the total variance attributable to differences between school in MATH, KNOW and APPLY drop significantly to 18% – 19%.

For REASON the variance between school drop to 14%. However, for SELFCONCEPT and SELFEFF, the variance between school was not significantly different.

The contribution of school and classroom climate to student learning is now better (after including school climate) compared to only including all student, teacher, and school background variables. The overall goodness of fit of the model (percentage of total variance explained) is improved, explaining approximately 25% - 26% of the total variance in MATH, KNOW, and APPLY; and 19% in REASON. For SELFCONCEPT and SELFEFF, the total variance explained considerably improved, explaining 13% and 22% respectively. That total variance explained, even not high but can indicate the influence of school and classroom climate on student achievement and self-belief. 25% total variance explained was acceptable since this model did not include students' prior achievement as one of the most significant predictors (Muijs & Reynolds, 2003; Muñoz-Chereau, 2013; Salim, 2011; Timmermans & Thomas, 2014). For comparison, Muñoz-Chereau (2013) that aimed to search fairer model of school effectiveness in Chile, found that by adding prior attainment, the goodness of fit improved dramatically compare to model that only includes student background variables, explaining 63% compared to 16% respectively.

Regarding school and classroom variance explained, this model explains 57% to 60% school variance in math and its cognitive domains, 70% in SELFEFF, and 46% in SELFCONCEPT. Interestingly, SELFCONCEPT has the highest classroom variance explained, approximately 65% (after controlling student, teacher, and student background variables). This high classroom variance explained (65%), which indicate the teacher's role in the classroom, can be explained by using the social learning theory of Bandura (Bandura, 1997; Bandura & Walters, 1963). In this situation, the teacher acts as a reference or model, and the student then makes the teacher a role model. Research by Cheng (2016) also supports this possibility. Cheng (2016) found that students' non-cognitive outcomes are developed by modelling their teacher.

For student's characteristics, gender has no significant relationship with student academic achievement but have a significant relationship with SELFCONCEPT and SELFEFF, where the boys have higher score in self-belief compared to the girls. This result is consistent with previous research such as Hergovich et al. (2004) found that girls' self-concept in math was lower compared to the boys. They added, girls' self-concept much depends on teachers' and parents 'judgements, whereas boys were not.

Next, for teacher characteristics, teacher's age found to have a significant relationship with student achievement. However, the relation was not linear, because student tends to have lower achievement with younger (less than 40 years old) and with older teachers (more than 60 years old). This result indicated that teachers within average age (40 -59) have experiences needed in managing their classroom. Teachers under '30s considerably have not enough experiences, and they are relatively young, so they may not understand their student' behaviour better. Contrary, teachers within 40 to 59 have more experiences.

For self-beliefs outcomes, teachers' age have insignificant factors, but teachers' major was a matter. The student with non-mathematics teachers tends to have higher SELFCONCEPT and SELFEFF.

Regarding school level characteristics or school context, average SES has a significant relationship with student achievement but not on student self-beliefs. This result is consistent with other research which mostly found that school with a low-SES group of students are often under-resourced and can affect student attainment (Aikens & Barbarin, 2008; Timmermans & Thomas, 2014). Also, other research found that the socioeconomic composition of student had a significant effect on student achievement (Muijs et al., 2010).

The other school-level factors, like school location and school types, now have no significant relationship with all the outcomes. This result is unexpected and exciting, because by taking into account school and classroom climate factors, the effect of school types mainly

being private *madrasah* become insignificant. This finding indicates that school and classroom climate factors are essential in reducing the negative impact of being private *madrasah*. Previously, most of the research found that *madrasah* have significantly lower achievement compared to other school types. (Ali et al., 2011; ADB, 2014; Ghozali, et al., 2013). Of the school and classroom climate factors that had a significant relationship with math achievement were student safety (not being bullied), school discipline and student engagement in a math lesson.

However, the relationship between student safety and math achievement was negative. This finding is different from previous research in the field that mostly found that a safe and orderly environment had a significant positive effect (e.g., Sammons et al., 1997). But this research also found that school discipline as the measure of the orderliness of the school had a significant correlation with student achievement. Therefore, this research was not entirely different from previous research. The correlation between school discipline and math achievement was more prominent compared to the correlation coefficient between student safety (not being bullied) with student math achievement.

Student engagement in math lesson was the highest coefficient correlation compared to other significant factors of school climate. It indicates that this factor was prominent in explaining variance in math achievement.

7.2.2 Implications of key findings RQ1

This study provides an initial descriptive approach to explore the potential influence of school climate on student learning outcomes in the Indonesian context. There are some evidence suggesting that there is room for improvement in creating a positive school climate. By taking into account school and classroom climate factors, the effect of school types particularly being private general school/*madrasah* become not significant. This finding indicates that once school and classroom climate factors are taken into account, the apparent

difference in student learning outcomes associated with attending private *madrasah* or private schools becomes non-significant in the Indonesian context. In other words, school climate rather than school type, perhaps a more pertinent key factor in promoting student learning outcomes. As mentioned in Chapter 1, this is important because the school climate is seen as a malleable aspect of education that school or local government can manipulate (Voight et al., 2013; Wang & Degol, 2016). So, it can be intervened immediately without waiting for a long time to change the curriculum or policy. Thus, this finding informs policymakers, teachers, and other stakeholders to have more focus on building a positive school and classroom climate to improve students' achievement irrespective of school types.

7.2.3 Discussion of key findings of RQ2: How do school stakeholders (headteachers, teachers, and students) from 4 different schools experience their respective school climate (headteacher, teacher, and student)?

7.2.3.1 Key finding 6: Similarities across all school

Applying DMEEs educational effectiveness evaluation dimensions (frequency, focus, stages, quality, and differentiation) in assessing school climate, was useful to describe the similarities and differences in the four selected case schools. Richer differences between school climate of a high and low performing school can be found as well as between different types of school in Indonesia context (general school and *madrasah*). This research, therefore, suggests that measuring/evaluating school climate without using effectiveness evaluation dimensions cannot differentiate clearly how school climate influences student learning outcomes. The similarities of school and classroom climate factors across all schools including three factors as proposed by Thapa et al. (2013). The three factors of school climate are: (1) safety (frequency, focus, stage, differentiation), (2) relationship (frequency, stage, differentiation), and (3) teaching and learning (frequency).

The safety factors of school climate were not different across schools in term of its frequency, focus, stage, and differentiation. This finding is quite similar to the finding in quantitative phase which also found that school safety factors (i.e., teacher safety – see Model 6) had no significant relationship with student achievement. Additionally, student emotional safety (not being bullied) had significant negative effect on student learning outcomes. The finding in this qualitative phase may strengthen that finding (in quantitative phase), because students in all schools, whether in low and high performing schools had similar perception about their school safety.

This finding is inconsistent with most of the empirical research which found school safety is one of school climate factor that influences student learning outcomes (Cohen et al., 2009; Thapa et al., 2013; Wang & Degol, 2016; Wang et al., 2014). However, the possible explanation is the student and other school member perception on what behaviours or actions represent bullying or safety as mentioned before (Lai et al., 2008; Widayanti & Siswati, 2009).

In term of relationship, indeed, some studies (i.e., Muñoz-Chereau, 2013) found that high performing schools typically have better school climate factors, in this case, is positive interpersonal relationships among staff. However, in the context of Indonesia, this result is relatively different. There were no differences between high and low performing school. This finding also consistent with phase 1 of this study (quantitative) result, which also found that there was no significant relationship between relationship factor (i.e., teacher-teacher relationship, see Model 6) with student learning outcomes, particularly in academic achievement of student. The explanation of this phenomenon can be related to Indonesian culture. As explained in Chapter 2 about Indonesian culture, it is argued that the Indonesian culture (i.e., harmony, collectivist minded) may constrain the working relationships of people in educational institutions, including schools (Dardjowidjojo, 2001). According to Hofstede (1986), the society with high power distance index tends to make everything harmonise, for

example students are not expected to challenge their teacher's views. Maintaining a good relationship is a must in this culture, even though it may not directly contribute to the results of academic performance.

7.2.3.2 Key Finding 7: Differences between high and low performing schools

As mention above, DMEE's educational effectiveness evaluation dimensions can make differences in how school and classroom climate are assessed. Therefore, the differences between high and low performing schools were mainly on its evaluation dimension rather than its school climate factors. This study found that most of the differences between high and low performing schools are on the *focus* and *quality* effectiveness evaluation dimension of DMEE.

In term of quality of safety, this finding is in line with most of empirical research which found school safety is one of school climate factor that influences student learning outcomes (Cohen et al., 2009; Thapa et al., 2013; Wang & Degol, 2016; Wang et al., 2014). School safety is required if academic achievement and self-development are set as the aims of schooling (Devine & Cohen, 2007).

In the teaching learning process, mainly happens in the classroom, the classroom serves as a fundamental position in which students and teachers create and manage a climate that influences all of the students within the group's academic emotional and social adaptation (Farmer et al., 2010; Kutsyuruba et al., 2015). Also, most of students may never reach the highest minimum standards and understand full potential until they experience a positive and supportive climate (Kutsyuruba et al., 2015; Urban, 1999). Loukas (2007) also argued that a safe school is the foundation of the healthy and stable school environment in which students succeed academically as well as emotionally and socially.

For institutional environment, also consistent with previous research on school climate which found that environmental surrounding plays a critical role to shape school member's experiences and related to academic learning outcomes (Cohen, 2006, 2009; Loukas, 2007;

Thapa et al., 2013; Wang & Degol, 2016). The physical environment includes: the suitability of the school physical environment, building management and facilities, and the availability and distribution of educational resources (Wang & Degol, 2015).

In term of school improvement process, the finding is also consistent with other research such as Fernandez (2011) found the relation between the school improvement plan and overall student performance in mathematics and reading is strong and consistent. Creemers (1999), who did research on primary school improvement in Indonesia, also argued the importance of school improvement. The research focused on teacher professional development, providing of textbooks, community involvement and school administration and management.

7.2.3.3 Key Finding 8: Differences between religious (*madrasah*) and non-religious (general) schools

Major differences between *madrasah* and general school mainly based on non-academic teaching and learning focuses. Madrasah's non-academic emphasis is primarily on the training of Islamic rituals. Madrasah, for example, supposed to train the students for Islamic funeral services.

Contrary, in general schools, regarding non-academic focus, the school did not have unique focus. Student in general schools only being taught formal non-academic activities as printed national curriculum, for example to develop student good character. This result is a reflection of Ma'zumi et al. (2012) view. He argued that madrasah is a school that focuses on transmitting Islamic values, to create young Muslims' role models. Tan (2014b) added, Islamic schools offer a variety of opportunities for students to improve their life skills that helps students internalise the principles and values they have studied and put them into practice. This differences also can create different culture between *madrasah* and general school. This finding also confirm that school and *madrasah* have different aims and focuses, and therefore can be use as basic information for students and parents to choose between the school type.

7.2.3.4 Key finding 9: New school climate factors that emerged from the data relevant to highlight potential factor of school climate to differentiate between the more and less performed schools

There are two new themes emerged from the qualitative data analysis process, namely the orientation of education and culture.

First, the education orientation influences all related stakeholders including teachers, headteachers, and policymakers in delivering education system, what is seen as an essential aspect of education, how to achieve it and how to do an evaluation of learning outcomes (Armstrong, 2006). These findings indicate that the learning results are classically determined by how well the student performs during the exam. (Berkhout et al., 2019; Effendi & Suyudi, 2016; Furaidah et al., 2015; Ragatz et al., 2015). This orientation was also reinforced on the next level of education by the school selection system which overlooks the score of the students in the national exam. Exam orientation leads the teaching learning process based solely on exam success and would have less emphasis on self-development and creativity of the students (Armstrong, 2006; Kirkpatrick & Zang, 2011).

Exam-oriented education, however, also does not only occur in the Indonesian context. Many countries such as the UK, the US, China, Tanzania and Kenya also have problems with this examination-oriented schooling (Jennings & Bearak, 2014; Mackatiani, 2017; Marshall, 2017; Salim, 2011). Exam-oriented education affects the teaching strategies and approaches used by teachers. Among the worst practices is when teachers only teach the students only to pass the test (Furaidah et al., 2015; Kirkpatrick & Zang, 2011). Consequently, it can adversely affect both critical and rational thinking in a negative way (Mackatiani, 2017; Revina, 2017).

Second, culture; culture is one of an essential factor of school climate in Tagiuri's model (Tagiuri, 1968). Tagiuri's model proposed a taxonomy of school climate, including ecology, milieu, social system, and culture. Therefore, culture is not a new factor in school

climate research, however, this research finding gives more emphasis that culture should be taken into account when evaluating the school climate, particularly in the Indonesia context. This research, therefore, following other researcher have argued that culture is part of the school climate (see Anderson, 1982; Tagiuri, 1968; Van Houtte, 2005)

The differences between *madrasahs* and general schools in this research are predominantly about the different culture. Culture is the set of stable social values that form beliefs and behaviours over time (Peterson & Deal, 2009). In fact, the differences in culture because these two types of schools applied different bases as well as being managed by the different ministry.

The student learns only two hours of Islamic Religious subject per week in general schools. In comparison, students need to study about five to six hours a week in *madrasah*s and as well as all subjects as taught in general schools. The *madrasah* student has additional classes, such as *Quran* and *Hadith* studies, Fiqh (jurisprudence), *Aqidah* (theology), *Akhlaq* (virtue), and Islamic History (Indonesia Ministry of Religious Affairs, 2014; Indonesia Ministry of Religious Affairs, 2014; Tan, 2014b), aimed primarily at transmitting religious principles, knowledge, and culture, so that principles can motivate and encourage the student to perform Islamic practices (Kholily, 2017).

7.2.4 Implications of key findings RQ2

First, the main implication of key finding in RQ2 is the usefulness of the effectiveness dimension in measuring school climate. By using the DMEE's effectiveness evaluation dimensions, the dynamic of the school climate can be better more systematically evaluated as explained in previous sections. Therefore, as implication, this research suggests applying effectiveness evaluation dimensions of DMEE (frequency, focus, quality, stage, and differentiation) when measure or evaluate school climate factors. This research also suggests that two influential evaluation dimensions in measuring school climate; focus and quality. This

is in line with Creemers and Kyriakides (2008) suggestion it was not possible to measure/evaluate the impact of one explanatory variable if only using frequency dimension. However, strongly argue to apply all the effectiveness evaluation dimensions.

Although this study found that applying the effectiveness evaluation dimensions in qualitative method is useful, however, this finding is tentative since most of SER that adopted DMEE and its evaluation dimensions were mostly use quantitative method (Creemers & Kyriakides, 2010b; Kyriakides & Creemers, 2009; Kyriakides et al., 2013; Panayiotou et al., 2016).

Second, in measuring school climate particularly in the Indonesian context and the country that have similar characteristics should consider two other school climate factors, namely the orientation of education and culture. This two school climate factors and its effectiveness evaluation dimensions could be used in the future as an evaluation list to identify and to capture the diversity of school climate. However, the use of those two factors is tentative and not final for measuring school climate, which is apparently beyond the scope of this study.

7.3 Overall conclusions: linking RQ1 and RQ2 findings and implications

7.3.1 Conclusion 1: School climate relationship with student outcome learning were explained in the Indonesian context.

Multilevel analysis followed by the illustrative case study providing evidence and a more comprehensive understanding of school climate in the Indonesian context. This can be utilised to explain school practices and how school climate can act as a protective factor to support better student outcomes, irrespective of school type. This result confirms the view that sees the school climate as a malleable aspect of education that school or local government can manipulate (Voight et al., 2013), without waiting for a long time to change the policy, the intervention can be conducted immediately.

However, the implemented design approach needs to be taken into account with caution, as the explanations provided in the study are not causal, particularly because the data used in the study is cross-sectional. Moreover, it is essential to remind that the qualitative findings may not be generalised to other schools. So, what was relevant in the four selected schools, may not apply to other schools, even in the same context.

This study linked two methodological approaches: quantitative, by using international large-scale assessment (TIMSS), and qualitative analysis of school and classroom climate. In doing so, school and classroom climate relationship and different types of outcomes of learning (academic and non-academic) were explained in the Indonesian context.

7.3.2 Conclusion 2: The use of effectiveness dimension of the dynamic model of educational effectiveness in assessing school and classroom climate were useful and give a more informative explanation.

The study provided original empirical evidence concerning the using of effectiveness dimension adopted from DMEE (Creemers & Kyriakides, 2008). By using the effectiveness dimension, a more informative of how school and classroom climate perceived by the school's member can be obtained. This can be done by evaluating the school climate factors using effectiveness evaluation dimension of DMEE. Thus, each of school climate factor has its effectiveness evaluation dimensions. However, not all of the effectiveness dimensions can be applied to assess school climate factors. The informativeness of the dimension makes a clear distinction between more and less performed school, particularly in the Indonesian context. Thus, this study provides a robust approach to assessing school and classroom climate.

7.3.3 Conclusion 3: Potential additional factors for evaluating school climate in the Indonesian context

This study adopted, extended, and combined two established theoretical frameworks, not only utilising it into a new context. By doing so, not only giving a new idea to measure

school climate but also producing additional factors of school climate that may help to understand the school and classroom climate, particularly in the Indonesian context. The additional school climate factors (culture and education orientation) could be used in the future as an assessment list.

The culture factor is not a new factor in school climate research, as mentioned in Chapter 6. However, this study has found that the use of culture factor is emphasised particularly in the Indonesian context which has dual curriculum systems to deliver its education system. These two systems have been identified as having different educational goals. *Madrasah* mainly aims to provide education to young Indonesian Muslim, so they can be better in the implementation of religious practices which based on Islamic values. Meanwhile, the general schools have different value which is based on Indonesian culture in general. Therefore, it is crucial to consider using this factor when assessing school climate particularly in the Indonesian context or other countries that have similar background.

Next, education orientation also needs to be assessed to look deeper on how a school runs its daily operations. It has been argued that the excessive emphasis on academic performance may lead education to only measure cognitive development of the student. Whereas the aim of education also to develop student social-emotional development. The factor of school climate (that has been adopted in this research) related to education orientation is teaching-learning term, which implicitly measures academic climate. By using education orientation factor, the researcher or educators can explicitly and directly measure their education orientation that has been run in a school (whether more focus on academic or development of student as a whole person).

7.4 Strengths and limitations of the study

7.4.1 Strengths of the study

7.4.1.1 Comprehensive data analysis

Using TIMSS datasets combined with qualitative analysis to study the relationship between Indonesian student performance and schools and classroom climate allowed getting a better understanding to explain school and classroom climate in the Indonesian context. This original approach, to the researcher knowledge, never been done in the previous study (i.e., Anderson, 1982; Wahyudi & Darrel, 2006; Carrasco Ogaz, 2016) particularly in the Indonesian context. This study strongly supports the value of mixed method approach where studies using quantitative techniques are strengthened if followed by a qualitative study to explore the processes that potentially cause observed effects (Thomas & Mortimore, 1996; Thrupp, 2001; Teddlie & Sammons, 2010).

7.4.1.2 Considering all types of schools present in Indonesia.

This research addressed the topic of school effectiveness in different types of schools in Indonesia (public general, public *madrasah*, private general, and private *madrasah*). This is significant because most of the studies carried out previously have not covered the full range of schools. By including all school types, this study not only echoing the two bodies of educational responsibility but also helps to explain the almost true educational reality in the Indonesian context. This characteristic of the sample analysed gave better explanation to the analysis.

7.4.1.3 Including a different range of educational outcomes: academic and self-belief

Most of the previous SER studies conducted have focused on academic achievement only. In the recent development of SER, many researchers have argued that giving attention to academic achievement exclusively will provide less evidence of school setting or the student non-academic outcomes (Creemers & Kyriakides, 2010; Knuver & Brandsma, 1993;

Mortimore, 1988; Thomas et al., 2000). As Guskey (2012) argued that student achievement is a multidimensional construct that can address different domains of learning, this research measured the broad range of educational outcomes.

7.4.1.4 Combining two different theoretical frameworks

This research combined two different conceptual frameworks. School climate and SER. Educational effectiveness model of Creemers and Kyriakides (2008) has a limited explanation of school climate. However, they provide more chance in how school climate is assessed by taking into account the five evaluation dimensions of educational effectiveness (frequency, focus, stage, quality, and differentiation). By using the effectiveness evaluation dimensions, it is possible to measure school climate better and more systematically. So, it may get a better understanding of the whole school climate.

On the other hand, Thapa et al. (2013) provide a better explanation of school climate factors. By combining these concepts, it will measure the effectiveness of school climate, not just the school climate. So, the whole picture of school climate might be captured, and this approach may give more information to the policymakers, headteachers, teachers, as well as students how to develop and evaluate school climate and may lead to improve it.

7.4.2 Limitations of the study

It is essential to consider some limitations when interpreting the research findings. There are some limitations of this study in design and methodology that indicate directions for future research.

7.4.2.1 Cross-sectional and correlational design

Large-scale surveys provide a wealth of useful data, but most of the analyses in this project were cross-sectional rather than longitudinal. Within the framework of a cross-sectional design, it is possible to demonstrate the strength and consistency of correlational relationships across samples, to control for other potentially confounding variables, and to show the

similarity of findings across measures and informants. However, a cross-sectional design cannot explore longitudinal pattern such as student relative “value-added” progress over time to establish causal effects (Goldstein & Thomas, 2008; Payne & Payne, 2004).

7.4.2.2 Retrospective nature of the data

This study used quantitative data from TIMSS 2011, the retrospective nature of the data needs to be highlighted as a limitation of this study. The condition may be changed when the qualitative phase took place in 2015. So, the key point is to understand the findings of this research as referring to one point in time, not as an overall account of schools' performance.

7.4.2.3 Using a self-report questionnaire

TIMSS data used in this study is heavily based on questionnaires. Questionnaires are retrospective and dependent on the memory of participants. In addition to memory difficulties, informants may have incomplete or inaccurate knowledge, and their reports may be skewed by personal biases, social desirability, or other motives (Allen, 2017). In the TIMSS context, the biases also come from a different understanding of the questions since TIMSS use the same framework to different country, and therefore may lead to cultural bias (Baird et al., 2011). For example, about this cultural bias, Pollitt and Ahmed (2001) showed significant differences between countries’ achievement in PISA that appear linked to culture and language bias.

7.4.2.4 Measurement error

The measurement of school climate in this study utilised the availability of items in TIMSS data. Therefore, the measure of school and classroom climate may not measure all the school climate factor correctly. Goldstein (2004) critiqued the way IEA and OECD in creating psychometric scales of the measurement as theoretically lack procedure. This study also confirms that there was some problem with items collected in TIMSS where some items were removed because it was a misfit with the standard of item goodness of fit (Linacre, 2011a) in term of multidimensionality and reliability.

In term of math achievement itself, as mentioned in Chapter 3, TIMSS measure in achievement may not reflect Indonesian student achievement related to Indonesia curriculum design, since TIMSS delivers similar test across countries.

7.4.2.5 Time restrictions on qualitative data collection

The qualitative data collection challenged time constraints due to the difficulties of collecting data across different regions and in a country different from the one where the researcher was residing, as well as due to limitations derived from the scope of a PhD research. The data collection was within two months. Additionally, the location of data collection was spread over cities, provinces as well as islands in Indonesia. It would have been better to have an individual interview with teachers and students rather than just a focus group to obtain more detailed qualitative responses.

7.4.2.6 Researcher as an outsider

Despite how detailed the descriptions presented in the case study phase attempted to be, they portrayed things mainly from an outsider perspective, as the researcher was not a member of the participating schools. Therefore, the participant may not be fully open to the questions asked by the researcher, because people's willingness to share information, what people say to others, is undoubtedly influenced by what they think of the researcher (Drever, 2003). When doing this research, the researcher also had to face a challenging situation as an outsider. For example, the researcher was refused to take an interview in one school because the researcher had not known anyone in the school.

7.5 Further research

7.5.1 Using longitudinal data or trend data analysis.

Unfortunately, the data analysed in this research was from only cross-sectional data based on TIMSS. It means that its results will require further confirmation by using a longitudinal dataset. So, further research can fully apply the dynamic model in researching

school climate as well as to examine the school climate change over a different period. Or at least, the next study should examine the trend on school climate perception on TIMSS to get more understanding of similarity over time by using multiple data sets provide by TIMSS. Alternatively, to test the consistency of school climate effect in one country, a new study that compares TIMSS and PISA on measuring school climate would be helpful. For example, research by Rutkowski and Rutkowski (2009) which tested the trends in TIMSS responses over time by using three TIMSS data set, including 1995, 1999, and 2003 data together. Another example is a student by Lenkeit and Caro (2014) which analysed data from 4 cycles of the PISA to measure and compare educational quality between countries.

7.5.2 Using an improved measurement of school and classroom climate factors

This research, particularly in quantitative phase, cannot be fully examined all the school climate factors as in the Thapa et al. (2013) framework. Most of the measurements acted as a proxy to measure school climate and depended heavily on the availability of the data in TIMSS. Therefore, there is a need to develop their own school climate measure, so the usefulness of the effectiveness dimension can be thoroughly tested in measuring school and classroom climate. Better measurement will lead to a better result, if the measurements are improper, then the findings will be doubtful (Goldstein, 1997). However, this research re-examined the reliability and validity of the scale measurement of school climate used using Rasch Model as mentioned in Chapter 4.

7.5.3 Using different academic achievement

To test the consistency of school climate effect, further research should consider including more diverse student outcome subjects such as literacy and science as well as mathematics. Some studies found differential effectiveness in a different school subject which might be linked to school climate. For example, Charalambous, et al. (2019) study on differential teacher effectiveness, found that teacher has a different effect on different subjects.

Charalambous and colleagues compared mathematics and physical education. Previously Ma (2001) found that student and school characteristics were differentially effective in various subject areas. The study examined differential effectiveness in mathematics, science, writing, and reading.

7.5.4 Looking more in-depth at the different aspect of differences between schools

This study focuses only on one aspect of school process in school effectiveness research. Therefore, the explanation of school process in different school type may be limited. In this research, total variance explained only 25-26 per cent for maths achievement and 13-22 per cent for self-beliefs. Thus, more than 70% variance explained may be due to other aspects of the school, classroom of student variables. To look more in-depth other aspects of school climate or other school process elements, the next study should use individual interview with teacher and student. This study found that the use of focus group interview is limited to enquire more in-depth opinions and perception of the school climate from the students and teachers. This is the limitation of using focus group interview, for example, some individuals dominated the groups, normative expressions and opinions, and the participant might not openly speak up about disagreements (Smithson, 2000).

7.6 Recommendations

The study findings suggest the following guidelines for the assessment and improvement of school climate in lower secondary schools in Indonesia.

1. This study found that school climate can act as a protective factor to mitigate the negative influence of school characteristic backgrounds (i.e., location, school type); therefore, the education authority can create positive school climate without waiting for curriculum or regulation change. It can immediately implement with the help and collaboration among all school stakeholders.

2. The study findings have indicated that student achievement and self-development in Indonesia lower secondary schools are related to what is happening inside the classroom and school in general. The study, therefore, recommends the improvement of pedagogical practice (improving the quality of classroom teaching and learning process) to create a positive classroom climate. The improvement can be delivered by schools and teachers.
3. The study found that the reasoning aspect of math achievement considerably lower compared to other cognitive domains (i.e., knowing and applying). Therefore, the education authority should improve teacher instructional strategy to enhance reasoning abilities by using various methods of teaching that can help the student to trigger this higher order thinking skills.
4. Since the study found that exam-oriented and teacher-centred learning are still practising, the education authority should seriously take action to evaluate this practice as well as to minimise school evaluation based on their performance in the national exam or other forms of the exam that excessively focus on the ability of student to pass exam.
5. The study found that in general, schools have less influence on promoting student self-development, which is against the Indonesian Regulation on the National Education System that aims to promote not only academic performance but also have to develop student personal development (Indonesia, 2003). Therefore, in the future education authority as well as teachers and schools should give more attention to student self-development, for example by also evaluate student personal development, asking the teacher to provide brief report on their student social-emotional skills (communication, self-beliefs, teamwork, etc).

6. The approach taken by the Indonesian government in improving the quality of education has been centralistic. This approach may not be effective considering that every school has its own school climate. An alternative approach is by implementing a decentralised education policy. A decentralisation policy provides schools with the ability to operate independently as per the needs of the community where the school is located. It allows for bottom-up interventions, in which school stakeholders (e.g., teachers, students, parents) work hand in hand to set their school goals, identify potential problems that could hinder the school from achieving its goals, and establish solutions to overcome the problems. Related to this, as an Indonesian academic, the author is obliged to conduct community services relevant to his expertise. For example, by giving free talks and lectures to teachers, students as well as the community around Medan, Indonesia, where the researcher resides. One way for the author to disseminate the present work is by presenting it on such community services sessions. The researcher will also write an article that summaries the findings of the current research in a national newspaper. This is important to reach a broader audience at the national level. Ultimately, the current research needs to be published in a reputable scientific journal. This approach will give stronger credence to the current project that appeals to academics and policymakers

7.7 Concluding statement

Finally, this study seeks to examine school climate, as one of key elements of school effectiveness research, and its relationship with both students' academic and affective outcomes in the Indonesian context. Although the findings are tentative and exploratory, this research is original in analysing and reporting on one specific aspect of the school process - school climate - to understand better the relationship between school climate and broader

student learning outcomes in Indonesia and other similar contexts and thereby contributing to the international school effectiveness knowledge base.

The study provided plausible explanations that have the potential to inform and enhance educational policy and practice. However, in terms of the findings of this study, it is important to acknowledge of Indonesian policymakers, practitioners and other educational authorities that the findings do not offer a plan or a recipe for improved school climate and should not be mechanically used without reference to the particular school setting. The findings are also retrospective, which refers to a group of students who have previously left the schooling system, and so the conditions for each school may not actually be a useful indicator of their future performance but would be useful for reflecting on past practice.

This study indeed reflects my professional experience as an academic in a university in Indonesia. This study has supported me to improve my experiences and expertise to explore different theoretical framework so that the effectiveness of school process can be evaluated in a better way. I have thus learned a new understanding that different research methods and paradigms can be integrated to produce a workable solution in a particular context, which strengthens my professional practice in using empirical data from multiple approaches.

References

Abu-Hilal, M., Abdelfattah, F., Alshumrani, S., Abduljabbar, A., & Marsh, H. (2013). Construct validity of self-concept in TIMSS's student background questionnaire: a test of separation and conflation of cognitive and affective dimensions of self-concept among Saudi eighth graders. *European Journal of Psychology of Education, 28*(4), 1201-1220. doi:10.1007/s10212-012-0162-1

Acharya, N., & Joshi, S. (2009). Influence of parents' education on achievement motivation of adolescents. *Indian Journal Social Science Researches, 6*(1), 72-79.

ADB. (2014). *Indonesia: Madrasah Education Development Project.* Retrieved from https://www.adb.org/documents/indonesia-madrasah-education-development-project

ADB. (2015). *Validation report: Indonesia: madrasah education development project*. Retrieved from http://www.adb.org/sites/default/files/PVR-383.pdf

Adel, T., & Zainal Ariffin, A. (2011). Principals' leadership style and school climate: teachers' perspectives from Malaysia. *International Journal of Leadership in Public Services, 7*(4), 314-333. doi:10.1108/17479881111194198

Afrianty, D., Hefner, R., & Azra, A. (2007). Pesantren and Madrasa Muslim Schools and National Ideals in Indonesia. In R. W. Hefner & M. Q. Zaman (Ed.), *Schooling Islam: the culture and politics of modern muslim education* (pp. 172-198). United States: Princeton University Press.

Aikens, N., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235. doi:10.1037/0022-0663.100.2.235

Al-Samarrai, S. (2013a). *Local governance and education performance: A survey of the quality of local education governance in 50 Indonesian districts*. Retrieved from Jakarta: https://openknowledge.worldbank.org/handle/10986/16765

Al-Samarrai, S. (2013b). *Tackling education inequality through better governance: Education in Indonesia.* Retrieved from http://blogs.worldbank.org/education/indonesia-tackling-education-inequality-through-better-governance

Ali, M., Kos, J., Lietz, P., Nugroho, D., Furqon, Zainul, A., & Emilia, E. (2011). *Quality of Education in Madrasah: Main Study*. Retrieved from Washington, DC: http://documents.worldbank.org/curated/en/2011/02/14048827/quality-education-madrasah-main-study

Allen, M. (2017). *The sage encyclopaedia of communication research methods* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc doi: 10.4135/9781483381411

Allison, P. (2001). *Missing data* (Vol. 136) California: Sage publications.

Anderson, A., Hamilton, R., & Hattie, J. (2004). Classroom climate and motivated behaviour in secondary schools. *Learning Environments Research, 7*(3), 211-225. doi:10.1007/s10984-004-3292-9

Anderson, C. (1982). The Search for School Climate: A Review of the Research. *Review of Educational Research, 52*(3), 368-420. doi:10.3102/00346543052003368

Anfara Jr, V., Brown, K., & Mangione, T. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher, 31*(7), 28-38.

Angus, L. (1993). The Sociology of School Effectiveness. *British Journal of Sociology of Education, 14*(3), 333-345. doi:10.1080/0142569930140309

Antecol, H., Eren, O., & Ozbeklik, S. (2015). The Effect of Teacher Gender on Student Achievement in Primary School. *Journal of Labor Economics, 33*, 63-89. doi:10.1086/677391

Armstrong, T. (2006). *The Best Schools: How Human Development Research Should Inform Educational Practice Association for Supervision and Curriculum Development.* Virginia: Association for Supervision and Curriculum Development.

Asadullah, M. (2018). Madrasah for girls and private school for boys? The determinants of school type choice in rural and urban Indonesia. *International Journal of Educational Development, 62*, 96-111.

Asparouhov, T. (2006). General multi-level modeling with sampling weights. *Communications in Statistics—Theory Methods, 35*(3), 439-460.

Astor, R., Benbenishty, R., & Estrada, J. (2009). School Violence and Theoretically Atypical Schools: The Principal's Centrality in Orchestrating Safe Schools. *American Educational Research Journal, 46*(2), 423-461. doi:10.3102/0002831208329598

Azra, A. (2003). *Surau, pendidikan Islam tradisional dalam transisi dan modernisasi* (1st ed). Ciputat: Logos Wacana Ilmu dan Pemikiran.

Azra, A. (2015). Genealogy of Indonesian islamic education: Roles in the modernization of muslim society. *Heritage of Nusantara: International Journal of Religious Literature, 4*(1), 85-114.

Azkiyah, S., Doolaard, S., Creemers, B., & Van Der Werf, M. (2014). The effects of two intervention programs on teaching quality and student achievement. *Journal of Classroom Interaction*, 4-11.

Bacharach, V., Baumeister, A., & Furr, R. (2003). Racial and Gender Science Achievement Gaps in Secondary Education. *The Journal of Genetic Psychology, 164*(1), 115-126. doi:10.1080/00221320309597507

Badowi, A. (2016). Ujian Nasional dan matinya kreativitas. *Media Indonesia*. Retrieved from https://mediaindonesia.com/read/detail/38105-ujian-nasional-dan-matinya-kreativitas

Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., & Daugherty, R. (2011). *Policy effects of PISA*. Oxford University Centre for Educational Assessment. https://ora.ox.ac.uk/objects/uuid:26c9fccd-ae47-424e-ba40-0c84ebedfc3e/download_file?file_format=pdf&safe_filename=Policy%2Beffects%2Bof%2BPISA.pdf&type_of_work=Report

Baker, M., Andriessen, J., & Järvelä, S. (2013). *Affective learning together : social and emotional dimensions of collaborative learning*. London: Routledge.

Bandura, A. (1997). *Self-efficacy : the exercise of control*. Basingstoke: W. H. Freeman.

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. London: Holt, Rinehart and Winston.

Barakat, B., & Bengtsson, S. (2018). What do we mean by school entry age? Conceptual ambiguity and its implications: the example of Indonesia. *Comparative Education, 54*(2), 203-224. doi:10.1080/03050068.2017.1360564

Barclay, J., & Wu, W. (1980). Classroom climates in Taiwanese and American elementary-schools - cross-cultural-study. *Contemporary Educational Psychology, 5*(1), 65-82. doi:10.1016/0361-476x(80)90027-2

Baskerville, R. (2003). Hofstede never studied culture. *Accounting, Organizations and Society, 28*(1), 1-14. doi:https://doi.org/10.1016/S0361-3682(01)00048-4

Baswedan, A. (2014, December). Gawat darurat pendidikan di Indonesia. In *The Emergency of Indonesian Education]. A paper delivered at the meeting between Ministry and Head of Education Offices Indonesia-wide in Jakarta, in December* (Vol. 1).

Bedi, A., & Garg, A. (2000). The effectiveness of private versus public schools: the case of Indonesia. *Journal of Development Economics, 61*(2), 463-494. doi:http://dx.doi.org/10.1016/S0304-3878(00)00065-1

BERA. (2011). *Ethical guidelines for educational research 2011*. BERA. https://www.bera.ac.uk/publication/bera-ethical-guidelines-for-educational-research-2011.

Berkhout, E., Pradhan, M., Rahmawati, Suryadarma, D., & Swarnata, A. (2019). *Cheating on national exams in Indoensia: How big is the problems*. Paper presented at the RISE Annual Conference, Banda Aceh, Indonesia.

Bernstein, B. (1970). Education cannot compensate for Society. *New Society*, 344-347.

Bertling, J., Marksteiner, T., & Kyllonen, P. (2016). General noncognitive outcomes. In *Assessing Contexts of Learning* (pp. 255-281). New York: Springer.

Best, M., Knight, P., Lietz, P., Lockwood, C., Nugroho, D., & Tobin, M. (2013). *The impact of national and international assessment programmes on education policy, particularly policies regarding resource allocation and teaching and learning practices in developing countries. Final report.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

Bhatla, N., Achyut, P., Khan, N., & Walia, S. (2014). *Are schools safe and gender equal spaces? Finding from a baseline study of school related gender-besed violence in five countries in Asia*.

Biggs, J. (1998). Assessment and classroom learning: a role for summative assessment?. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 103-110.

Billig, M. (2013). Learn to write badly: How to succeed in the social sciences. *Cambridge, UK: Cambridge*.

Bima, L., & Yusrina, A. (2018). Africa: The Conversation. Retrieved from: http://the conversation.com/more-prosperous-teachers-have-noimpact-in-the-quality-of-education

Bjork, C. (2004). Decentralisation In Education, Institutional Culture And Teacher Autonomy In Indonesia. *International Review of Education, 50*(3), 245-262. doi:10.1007/s11159-004-2622-6

Bjork, C. (2005). *Indonesian education: Teachers, schools, and central bureaucracy*. London: Routledge.

Bloom, B. (1976). *Human characteristics and school learning*. New York: London : McGraw-Hill.

Boikhutso, K. (2010). The theory into practice dilemma: Lesson planning challenges facing Botswana student-teachers. *Improving Schools*, *13*(3), 205-220.

Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences, 2nd ed*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Bong, M., & Clark, R. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist, 34*(3), 139-153. doi:10.1207/s15326985ep3403_1

Bong, M., & Skaalvik, E. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review*, *15*(1), 1-40. doi:10.1023/A:1021302408382

Boone, W., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education, 4*(1), 1416898. doi:10.1080/2331186X.2017.1416898

BPS. (2012). *Statistik Indonesia.* Jakarta: BPS

BPS. (2018). *Statistik Indonesia dalam infografis.* Jakarta: BPS

Bradshaw, C., Waasdorp, T., Debnam, K., & Johnson, S. (2014). Measuring school climate in high schools: a focus on safety, engagement, and the environment. *Journal of School Health, 84*(9), 593-604. doi:10.1111/josh.12186

Brault, M., Janosz, M., & Archambault, I. (2014). Effects of school composition and school climate on teacher expectations of students: A multilevel analysis. *Teaching and Teacher Education, 44*(0), 148-159. doi:http://dx.doi.org/10.1016/j.tate.2014.08.008

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Brincks, A., Enders, C., Llabre, M., Bulotsky-Shearer, R., Prado, G., & Feaster, D. (2017). Centering Predictor Variables in Three-Level Contextual Models. *Multivariate Behavioral Research, 52*(2), 149-163. doi:10.1080/00273171.2016.1256753

Brookhart, S. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice, 8*(2), 153-169.

Brookover, W., Schweitzer, J., Schneider, J., Beady, C., Flood, P., & Wisenbaker, J. (1978). Elementary School Social Climate and School Achievement. *American Educational Research Journal, 15*(2), 301-318. doi:10.2307/1162468

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The Utility of Template Analysis in Qualitative Psychology Research. *Qualitative Research in Psychology, 12*(2), 202-222. doi:10.1080/14780887.2014.955224

Brown, G., Brown, G. D., Brown, G. R., Gillian, B., & Yule, G. (1983). *Discourse analysis.* Cambridge: Cambridge university press.

Brunner, M., Krauss, S., & Kunter, M. (2008). Gender differences in mathematics: Does the story need to be rewritten? *Intelligence, 36*(5), 403-421. doi:https://doi.org/10.1016/j.intell.2007.11.002

Bryner, K. (2013). *Piety projects: Islamic schools for Indonesia's urban middle class.* (PhD Doctoral thesis). Columbia University,

Buchori, M. (2001). *Notes on education in Indonesia*. Jakarta Post

Burger, R. (2011). School effectiveness in Zambia: The origins of differences between rural and urban outcomes. *Development Southern Africa, 28*(2), 157-176. doi:10.1080/0376835X.2011.570064

Burns, R. (1982). *Self-concept development and education*. London: Holt, Rinehart and Winston.

Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. (2010). Organizing schools for improvement: Lessons from Chicago: University of Chicago Press.

Byrne, B. (1984). The General/Academic Self-Concept Nomological Network: A Review of Construct Validation Research. *Review of Educational Research, 54*(3), 427-456. doi:10.2307/1170455

Cai, J. (2017). Rasch Analysis of the Mathematics Self Concept Questionnaire. In. New York, USA: Cornell University.

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., & Hiebert, J. (2017). Making classroom implementation an integral part of research. *Journal for Research in Mathematics Education, 48*(4), 342-347.

Calsyn, R., & Kenny, D. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement. *Journal of Educational Psychology, 69*(2), 136-145. doi:10.1037/0022-0663.69.2.136

Carle, A. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC medical research methodology, 9*(1), 49.

Carrasco Ogaz, D. A. (2016). *Multivariate approaches to school climate factors and school outcomes.* (PhD Doctoral Thesis). University of Sussex, Sussex, United Kingdom.

Cartwright, F., Lalancette, D., Mussio, J., & Xing, D. (2003). *Linking provincial student assessments with national and international assessments*. Ottawa: Statistics Canada, Culture, Tourism & the Centre for Education Statistics.

Cavanagh, R., & Waugh, R. (2011). *Applications of Rasch Measurement in Learning Environments Research*. Boston: SensePublishers.

Chang, M. C., Shaeffer, S., Al-Samarrai, S., Ragatz, A., de Ree, J., & Stevenson, R. (2014). *Teacher Reform in Indonesia: The Role of Politics and Evidence in Policy Making*. Washington, DC: World Bank.

Chapman, C., Armstrong, P., Harris, A., Reynolds, D., & Sammons, P. (2012). School effectiveness an school improvement: research, policy, and practice. In P. Armstrong, A. Harris, D. Reynolds, & P. Sammons (Eds.), *School effectiveness an school improvement: research, policy, and practice*. Canada: Routledge.

Charalambous, C., Kyriakides, E., Kyriakides, L., & Tsangaridou, N. (2019). Are teachers consistently effective across subject matters? Revisiting the issue of differential teacher effectiveness. *School Effectiveness and School Improvement*, 30(4) . 353-379. doi:10.1080/09243453.2019.1618877

Chen, S., Lin, C., Wang, J., Lin, S., & Kao, H. (2012). A Cross-grade Comparison to Examine the Context Effect on the Relationships Among Family Resources, School Climate, Learning Participation, Science Attitude, and Science Achievement Based on TIMSS 2003 in Taiwan. *International Journal of Science Education*, *34*(14), 2089-2106. doi:10.1080/09500693.2012.701352

Chen, S., Yeh, Y., Hwang, F., & Lin, S. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences, 23*, 172-178. doi:https://doi.org/10.1016/j.lindif.2012.07.021

Cheng, A. (2016). *Teachers and the Development of Student Noncognitive Skills.* (PhD Thesis). University of Arkansas, Fayetteville, Arkansas.

Cheng, C. (1994). Classroom environment and student affective performance - an effective profile. *Journal of Experimental Education, 62*(3), 221-239.

Coates, D. (2003). Education production functions using instructional time as an input. Education Economics, 11, 273–292.

Codding, R., & Smyth, C. (2008). Using performance feedback to decrease classroom transition time and examine collateral effects on academic engagement. *Journal of Educational and Psychological Consultation, 18*(4), 325-345.

Coe, R., & Fitz-Gibbon, C. (1998). School Effectiveness Research: criticisms and recommendations. *Oxford Review of Education, 24*(4), 421-438. doi:10.1080/0305498980240401

Cohen, J. (2006). Social, Emotional, Ethical, and Academic Education: Creating a Climate for Learning, Participation in Democracy, and Well-Being. *Harvard Educational review, 76*(2). 201-237

Cohen, J. (2009). Transforming school climate: Educational and psychoanalytic perspectives: Introduction. *Schools. 6*(1), 99-103. doi:10.1086/597659

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School Climate: Research, Policy, Practice, and Teacher Education. *Teachers College Record, 111*(1), 180-213.

Coleman, J, Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity* (1066-5684). https://doi.org/10.1080/0020486680060504

Coopersmith, S. (1967). *The antecedents of self-esteem.* San Francisco: Freeman.

Cornell, D., Huang, F., Konold, T., Malone, M., Datta, P., Stohlman, S., & Meyer, J. P. (2016). *Development of a standard model for school climate and safety assessment: Final report.* Huntsville: National Children Advocacy Center.

Côté, J. (2009). Identity formation and self-development in adolescence. In *Handbook of adolescent psychology: Individual bases of adolescent development, Vol. 1, 3rd ed.* (pp. 266-304). Hoboken: John Wiley & Sons Inc.

Creemers, B. (1994). *The effective classroom*. London: Cassell.

Creemers, B. (1999). Educational effectiveness and improvement in developing societies: some experiences from the primary education quality improvement project in Indonesia. *Tertium Comparationis, 5*(1), 32-51.

Creemers, B., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement, 17*(3), 347-366. doi:10.1080/09243450600697242

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness : a contribution to policy, practice and theory in contemporary schools / Bert P.M. Creemers and Leonidas Kyriakides*. London : Routledge, 2008.

Creemers, B., & Kyriakides, L. (2009). Situational effects of the school factors included in the dynamic model of educational effectiveness. *South African Journal of Education*, *29*(3).

Creemers, B., & Kyriakides, L. (2010a). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409-427.

Creemers, B., & Kyriakides, L. (2010b). School Factors Explaining Achievement on Cognitive and Affective Outcomes: Establishing a Dynamic Model of Educational

Effectiveness. *Scandinavian Journal of Educational Research, 54*(3), 263-294. doi:10.1080/00313831003764529

Creemers, B., & Reezigt, G. J. (1999). The Role of School and Classroom Climate in Elementary School Learning Environments. In J. Freiberg (Ed.), *School Climate: Measuring, Improving and Sustaining Healthy Learning Environments* London: Falmer Press.

Creswell, J. (1994). *Research design : qualitative & quantitative approaches*. Thousand Oaks, Calif. ; London: Sage Publications.

Creswell, J. (2003). *Research design : qualitative, quantitative, and mixed methods approaches* (2nd ed.). London: Sage Publications.

Creswell, J. (2009). *Research design : qualitative, quantitative, and mixed methods approaches* (3rd ed.). London: Sage Publications.

Creswell, J. (2013). *Qualitative inquiry & research design : choosing among five approaches* (3rd ed.). London: Sage Publications.

Creswell, J.. (2014). *Research design : qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles: Sage Publications.

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). London: SAGE.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. doi:10.1037/h0040957

Croninger, R., Rice, J., Rathbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*, *26*(3), 312-324. doi:https://doi.org/10.1016/j.econedurev.2005.05.008

Crotty, M. (1998). *The foundations of social research*. London: Sage

Dallimore, E., Hertenstein, J., & Platt, M. (2004). Classroom participation and discussion effectiveness: student-generated strategies. *Communication Education, 53*(1), null-null. doi:10.1080/0363452032000135805

Damanik, E., & Aldridge, J. (2017). Transformational Leadership and its Impact on School Climate and Teachers' Self-Efficacy in Indonesian High Schools. *Journal of School Leadership, 27*(2), 269-296. doi:10.1177/105268461702700205

Dardjowidjojo, S. (2001). Cultural constraints in the implementation of learner autonomy: The case in Indonesia. *Journal of Southeast Asian Education, 2*(2), 309-322.

Darling-Hammond, L. (2000a). How Teacher Education Matters. *Journal of Teacher Education, 51*(3), 166-173. doi:10.1177/0022487100051003002

Darling-Hammond, L. (2000b). Teacher quality and student achievement. *Education policy analysis archives, 8*, 1.

Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education—A systematic literature review. *Thinking Skills and Creativity, 8*, 80-91. doi:https://doi.org/10.1016/j.tsc.2012.07.004

Dazeva, V., & Tarmidi, T. (2012). Perbedaan Kecerdasan Emosional Siswa Ditinjau dari Jenis Kegiatan Extrakurikuler. *Psikologia: Jurnal Pemikiran dan Penelitian Psikologi, 7*(2).

Deemer, S. (2004). Classroom goal orientation in high school classrooms: Revealing links between teacher beliefs and classroom environments. *Educational research, 46*(1), 73-90.

De Fraine, B., Van Damme, J., & Onghena, P. (2002). Accountability of Schools and Teachers: What Should Be Taken into Account? *European Educational Research Journal, 1*(3), 403-428. doi:10.2304/eerj.2002.1.3.2

De Ree, J., Muralidharan, K., Pradhan, M., & Rogers, H. (2017). Double for nothing? experimental evidence on an unconditional teacher salary increase in Indonesia. *Quarterly Journal of Economics.* 1-72

Denscombe, M. (2010). *Ground rules for social research : guidelines for good practice* (2nd ed.). Maidenhead: McGraw-Hill/Open University Press.

Denzin, N., & Lincoln, Y. (2005). *The SAGE Handbook of Qualitative Research.* London: Sage Publications.

Devine, J.., & Cohen, J. (2007). *Making your school safe: Strategies to protect children and promote learning*: Teachers College Pr.

Dewey, J. (1902). Interpretation of savage mind. *Psychological Review*, *9*(3), 217.

Dong, Y., & Peng, C. (2013). Principled missing data methods for researchers. *Springer Plus*, *2*(1), 222. doi:10.1186/2193-1801-2-222

Dorina, R. (2013). School climate as an important component in school effectiveness. *Academicus International Scientific Journal, 8*, 110-125. doi:10.7336/academicus.2013.08.06

Dorling, D., & Tomlinson, S. (2016). The creation of inequality: Myths of potential and ability. *Journal of Critical Education Policy Studies, 14*(3). 1-23.

Drever, E. (2003). *Using semi-structured interviews in small-scale research : a teacher's guide*. Glasgow: Scottish Council for Research in Education.

Dumaresq, R., & Blust, R. (1981). *School Climate Improvement: A Model for Effective School Change*. Harrisburg: Pennsylvania Department of Education

Earl, L. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks: Corwin Press.

Eccles, J. (2005). Influences of parents' education on their children's educational attainments: The role of parent and child perceptions. *London review of education, 3*(3), 191-204.

Eccles, J., Arberton, A., Buchanan, C., Janis, J., Flanagan, C., & Harold, R. (1993a). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choices. *Developmental perspectives on motivation, 40,* 145-208.

Eccles, J., Wigfield, A., Midgley, C., Reuman, D., Iver, D., & Feldlaufer, H. (1993b). Negative Effects of Traditional Middle Schools on Students' Motivation. *The Elementary School Journal, 93*(5), 553-574. doi:10.2307/1001828

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1). 15-24

Effendi, T., & Suyudi, I. (2016). *The Impacts of English National Examination in Indonesia.* Paper presented at the Ninth International Conference on Applied Linguistics (CONAPLIN 9).

Elias, M. (2003). *Academic and social-emotional learning*. Brussels, Geneva: International Academy of Education.

Eliot, M., Cornell, D., Gregory, A., & Fan, X. (2010). Supportive school climate and student willingness to seek help for bullying and threats of violence. *Journal of School Psychology, 48*(6), 533-553. doi:http://dx.doi.org/10.1016/j.jsp.2010.07.001

Elliott, J. (1994). The Teacher's Role in Curriculum Development: an unresolved issue in English attempts at curriculum reform. *Curriculum Studies, 2*(1), 43-69.

Elliott, J. (1996). School Effectiveness Research and its Critics: alternative visions of schooling. *Cambridge Journal of Education, 26*(2), 199-224. doi:10.1080/0305764960260205

Enders, C., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121-138. doi:10.1037/1082-989X.12.2.121

Eshach, H., Dor-Ziderman, Y., & Yefroimsky, Y. (2014). Question asking in the science classroom: Teacher attitudes and practices. *Journal of Science Education and Technology, 23*(1), 67-81.

Faisal, & Martin, S. N. (2019). Science education in Indonesia: past, present, and future. *Asia-Pacific Science Education, 5*(1), 4. doi:10.1186/s41029-019-0032-0

Faour, M. (2012). *The Arab World's Education Report Card School Climate and Citizenship Skills*. Washington DC: Carnegie Middle East Centre

Faridi, A., Bahri, S., & Nurmasitah, S. (2016). The problems of applying student centered syllabus of English in vocational high schools in Kendal regency. *English Language Teaching, 9*(8). 231-240

Farmer, T., Hamm, J., Petrin, R., Robertson, D., Murray, R., Meece, J., & Brooks, D. (2010). Supporting early adolescent learning and social strengths: Promoting productive contexts for students at-risk for EBD during the transition to middle school. *Exceptionality, 18*(2), 94-106.

Felner, R., Brand, S., DuBois, D., Adan, A., Mulhall, P., & Evans, E. (1995). Socioeconomic Disadvantage, Proximal Environmental Experiences, and Socioemotional and Academic Adjustment in Early Adolescence: Investigation of a Mediated Effects Model. *Child Development, 66*(3), 774-792.

Fenanlampir, A., Batlolona, J., & Imelda, I. (2019). Indonesian students in the context of TIMSS and PISA has not ended. *International Journal of Civil Engineering and Technology, 10*(2), 393-406.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods, 5*(1), 80-92. doi:10.1177/160940690600500107

Fernandez, K. (2011). Evaluating school improvement plans and their affect on academic performance. *Educational Policy, 25*(2), 338-367.

Field, A. (2005). *Discovering statistics using SPSS : (and sex, drugs and rock 'n' roll)* (2nd ed.). London: SAGE.

Filmer, D., & Pritchett, L. (1999). The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries. *Population and Development Review, 25*(1), 85-120.

Finlayson, D. (1987). School climate: an outmoded metaphor? *Journal of Curriculum Studies, 19*(2), 163-173. doi:10.1080/0022027870190205

Fisher, D., & Fraser, B. (1983). *Use of Classroom Environment Scale in Investigating Effects of Psychosocial Milieu on Science Students' Outcomes.* Paper presented at Annual Meeting of National Association for Research in Science Teaching, Dallas.

Fisher, D., & Fraser, B. (1990). *Validity and Use of the School-Level Environment Questionnaire.* Paper presented at Annual Meeting of American Educational Research Association, Boston.

Forshee, J. (2006). *Culture and Customs of Indonesia.* Westport: Greenwood Press.

Foy, P., Arora, A., & Stanco, G. (2013). *TIMSS 2011 User Guide for the International Database* Boston, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College andInternational Association for the Evaluation of Educational Achievement (IEA).

Foy, P., Brossman, B., & Galia, J. (2012). Scaling the TIMSS and PIRLS 2011 Achievement Data. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Fraser, B., Docker, J., & Fisher, D. (1988). Assessing and improving school climate. *Evaluation & Research in Education, 2*(3), 109-122. doi:10.1080/09500798809533248

Fraser, B., & Fisher, D. (1986). Using short forms of classroom climate instruments to assess and improve classroom psychosocial environment. *23*(5), 387-413. doi:10.1002/tea.3660230503

Freiberg, H. (1999). *School climate : measuring, improving and sustaining healthy learning environments*. London ; Philadelphia: Falmer Press.

Freiberg, H., & Stein, T. (1999). Measuring, improving and sustaining healthy learning envirnoments. In H. J. Freiberg (Ed.), *School Climate: Measuring, Improving and Sustaining Healthy Learning Environments*. Philadelphia, PA: Falmer Press.

French, D., Pidada, S., & Victor, A. (2005). Friendships of Indonesian and United States youth. *International Journal of Behavioral Development, 29*(4), 304-313.

Fung, F., Tan, C., & Chen, G. (2018). Student engagement and mathematics achievement: Unraveling main and interactive effects. *Psychology in the Schools, 55*(7), 815-831. doi:10.1002/pits.22139

Furaidah, F., Saukah, A., & Widiati, U. (2015). Washback of english national examination in the indonesian context. *TEFLIN Journal; Vol 26, No 1 (2015)*. doi:10.15639/teflinjournal.v26i1/36-58

Galecki, A., & West, B. (2013). Software for fitting multilevel models. In M. A. Scott, J. S. Simonoff, & B. D. Marx (Eds.), *The SAGE handbook of multilevel modeling*. Los Angeles: SAGE Publication Ltd.

Gardner, H. (2006). *The development and education of the mind : the selected works of Howard Gardner*. London: Routledge.

Geertz, C. (1973). *The interpretation of cultures: Selected essays by Clifford Geertz* (Vol. 5019). New York, USA: Basic books.

Gent, B. (2012). Inside madrassas: understanding and engaging with British-Muslim faith supplementary schools. *British Journal of Religious Education, 34*(3), 357-358. doi:10.1080/01416200.2012.707844

German, J., Villapando, K., Resilva, J., Quiambao, J., & Guevarra, A. (2020, April). Effects of Various Seating Arrangements on Academic Performance of Grade 11 Students in Statistics. In *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)* (pp. 953-957). IEEE.

Ghozali, A., Mudjahid, A. K., & Hayati, M. (2013). Madrasah education financing study: the education sector analytical and capacity development partnership [Press release]

Gibson, N., & Olejnik, S. (2003). Treatment Of Missing Data At The Second Level Of Hierarchical Linear Models. *http://dx.doi.org/10.1177/0013164402250987*. doi:10.1177_0013164402250987.

Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journa , 204*(6), 291-295. doi:10.1038/bdj.2008.192

Given, L. & Saumure, K. (2008). *Trustworthiness. The Sage Encyclopedia of Qualitative Research Methods. SAGE Publications, Inc*. Thousand Oaks: SAGE Publications, Inc.

Globe Project. (2020). globeproject.com

Glover, D., & Coleman, M. (2005). School culture, climate and ethos: interchangeable or distinctive concepts? *Journal of In-Service Education, 31*(2), 251-272. doi:10.1080/13674580500200278

Goldstein, H. (1979). Consequences of using the Rasch model for educational assessment. *British Educational Research Journal, 5*(2), 211-220.

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement, 8*(4), 369-395. doi:10.1080/0924345970080401

Goldstein, H. (2004). International comparisons of student attainment: some issues arising from the PISA study. *Assessment in Education: Principles, Policy & Practice, 11*(3), 319-330. doi:10.1080/0969594042000304618

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). Chichester: Wiley.

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D., & Thomas, S. (1993). A Multilevel Analysis of School Examination Results [1]. *Oxford Review of Education, 19*(4), 425-433. doi:10.1080/0305498930190401

Goldstein, H., & Thomas, S. M. (2008). Reflections on the international comparative surveys debate. *Assessment in Education: Principles, Policy & Practice, 15*(3), 215-222. doi:10.1080/09695940802417368

Goldstein, H., & Woodhouse, G. (2000). School effectiveness research and educational policy. *Oxford Review of Education, 26*(3-4), 353-363. doi:Doi 10.1080/713688547

Gooding, Y. (2001). *The relationship between parental educational level and academic success of college freshmen.* (PhD thesis). Iowa State University, Ames, Iowa.

Goossens, N., Camp, G., Verkoeijen, P., Tabbers, H., & Zwaan, R. (2014). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition, 3*(3), 177-182.

Graham, J. (2008). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology, 60*(1), 549-576. doi:10.1146/annurev.psych.58.110405.085530

Grandy, G. (2010). Instrumental case study. In A. J. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 474-475). Thousand Oaks, California.

Gray, J. (1996). *Merging traditions : the future of research on school effectiveness and school improvement*. London: Cassell.

Gray, J. (2004). School effectiveness and the ‘other outcomes’ of secondary schooling: a reassessment of three decades of British research. *Improving Schools, 7*(2), 185-198. doi:10.1177/1365480204047348

Gray, J., Goldstein, H., & Thomas, S. (2001). Predicting the future: the role of past performance in determining trends in institutional effectiveness at A level. *British Educational Research Journal, 27*(4), 391-405. doi:Doi 10.1080/01411920120071425

Greene, B., Miller, R., Crowson, H., Duke, B., & Akey, K. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology*, *29*(4), 462-482. doi:10.1016/j.cedpsych.2004.01.006

Greene, J., & Caracelli, V. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation, 1997*(74), 5-17. doi:10.1002/ev.1068

Gremmen, M., van den Berg, Y., Segers, E., & Cillessen, A. (2016). Considerations for classroom seating arrangements and the role of teacher characteristics and beliefs. *Social Psychology of Education*, *19*(4), 749-774.

Grotevant, H. (1998). *Adolescent development in family contexts.* In W. Damon & N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development* (p. 1097–1149). John Wiley & Sons, Inc.

Guba, E., & Lincoln, Y. (1994). Competing paradigms in qualitative research. In *Handbook of qualitative research.* (pp. 105-117). Thousand Oaks: Sage Publications, Inc.

Guinness, P. (1986). *Harmony and hierarchy in a Javanese kampung*: Oxford: Oxford University Press.

Gür, B., Çelik, Z., & Özoğlu, M. (2012). Policy options for Turkey: a critique of the interpretation and utilization of PISA results in Turkey. *Journal of Education Policy, 27*(1), 1-21. doi:10.1080/02680939.2011.595509

Guskey, T. (2012). Defining Student Achievement. In J. Hattie & E. M. Anderman (Eds.), *International Guide to Student Achievement*. New York: Routledge.

Gutman, L., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Endowment Foundation*, *59*(22.2).

Haack, S. (1976). The pragmatist theory of truth. *The British Journal for the Philosophy of Science, 27*(3), 231-249.

Hahs-Vaughn, D. (2005). A Primer for Using and Understanding Weights With National Datasets. *The Journal of Experimental Education, 73*(3), 221-248. doi:10.3200/JEXE.73.3.221-248

Hair Jr, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate Data Analysis* (7th Ed.). England: Pearson Education Limited.

Halpin, A., & Croft, D. (1963). *The organizational climate of School*. Chicago: Midwest Administration Center of University of Chicago.

Hammersley, M., & Atkinson, P. (2007). *Ethnography : principles in practice* (3rd ed.). Abingdon: Routledge.

Hasan, N., & Jihad, L. (2008). The Salafi Madrasas of Indonesia. *The Madrasa in Asia: Political Activism and Transnational Linkages*, 247-274.

Hesselmann, G. (2018). Applying linear mixed effects models (LMMs) in within-participant designs with subjective trial-based assessments of awareness—a caveat. *Frontiers in Psychology*, *9.* 788. 1-5.

Hastings, N., & Schwieso, J. (1995). Tasks and tables: The effects of seating arrangements on task engagement in primary classrooms. *Educational Research, 37*(3), 279-291.

Hattie, J. (1985). Methodology Review: Assessing Unidimensionality of Tests and ltenls. *Applied Psychological Measurement, 9*(2), 139-164. doi:10.1177/014662168500900204

Hattie, J. (2012). *Visible learning for teachers : maximizing impact on learning*. Abingdon: Routledge.

Heck, R. (2009). Teacher effectiveness and student achievement: Investigating a multilevel cross-classified model. *Journal of Educational Administration, 47*(2), 227-249.

Hendajany, N. (2016). The Effectiveness of Public Vs Private Schools in Indonesia. *Journal of Indonesian Applied Economics, 6*(1), 66-89.

Hergovich, A., Sirsch, U., & Felinger, M. (2004). Gender Differences in the Self-Concept of Preadolescent Children. *School Psychology International, 25*(2), 207-222. doi:10.1177/0143034304043688

Hiebert, J., & Stigler, J. (2000). A Proposal for Improving Classroom Teaching: Lessons from the TIMSS Video Study. *101*(1), 3-20. doi:10.1086/499656

Hill, P., & Rowe, K. (1996). Multilevel Modelling in School Effectiveness Research. *School Effectiveness and School Improvement, 7*(1), 1-34. doi:10.1080/0924345960070101

Hinkin, T., Holtom, B. C., & Klag, M. (2007). Collaborative Research. *Organizational Dynamics, 36*(1), 105-118. doi:10.1016/j.orgdyn.2006.12.005

Hofstede, G. (1984). The cultural relativity of the quality of life concept. *The Academy of Management Review, 9*(3), 389-398. doi:10.2307/258280

Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations, 10*(3), 301-320. doi: 10.1016/0147-1767(86)90015-5

Hofstede, G. (1993). Cultural Constraints in Management Theories. *The Executive, 7*(1), 81-94.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. . *Online Readings in Psychology and Culture, 2*(1). doi:10.9707/2307-0919.1014

Hofstede, G., Hofstede, G., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind*, (3rd ed). New York: McGraw-Hill Education.

Holloway, I., & Todres, L. (2003). The Status of Method: Flexibility, Consistency and Coherence. *Qualitative Research, 3*(3), 345-357. doi:10.1177/1468794103033004

Holth, P. (2001). The Persistence of Category Mistakes in Psychology. *Behavior and Philosophy, 29*, 203-219.

House, R., Dorfman, P., Javidan, M., Hanges, P., & de Luque, M. (2013). *Strategic Leadership Across Cultures: GLOBE Study of CEO Leadership Behavior and Effectiveness in 24 Countries*. London: Sage Publications.

Howe, K. (1988). Against the Quantitative-Qualitative Incompatibility Thesis or Dogmas Die Hard. *Educational Researcher, 17*(8), 10-16. doi:10.3102/0013189x017008010

Hox, J. (2002). *Multilevel analysis : techniques and applications*. Mahwah.: Lawrence Erlbaum Publishers.

Hoy, W. (1990). Organizational Climate and Culture: A Conceptual Analysis of the School Workplace. *Journal of Educational and Psychological Consultation, 1*(2), 149-168. doi:10.1207/s1532768xjepc0102

Hoy, W. (2012). School characteristics that make a difference for the achievement of all students. *Journal of Educational Administration, 50*(1), 76-97. doi:10.1108/09578231211196078

Hoy, W., & Hannum, J. (1997). Middle School Climate: An Empirical Assessment of Organizational Health and Student Achievement. *Educational Administration Quarterly, 33*(3), 290-311. doi:10.1177/0013161X97033003003

Hoy, W., & Miskel, C. (1996). *Educational administration : theory, research, and practice* (5th ed.). New York: McGraw-Hill.

Hoy, W., & Miskel, C. (2005). *Educational administration : theory, research, and practice* (7th ed.). Boston: McGraw-Hill.

Hoy, W., Smith, P., & Sweetland, S. (2002). The development of the organizational climate index for high schools: Its measure and relationship to faculty trust. *The High School Journal, 86*(2), 38-49.

Hoy, W., & Tarter, C. (1997). *The road to open and healthy schools: A handbook for change*. California: Corwin Press Thousand Oaks.

Hoy, W., Tarter, C., & Hoy, A. (2006). Academic Optimism of Schools: A Force for Student Achievement. *American Educational Research Journal, 43*(3), 425-446. doi:10.2307/4121765

Hoy, W., Tarter, C., & Kottkamp, R. (1991). *Open schools, healthy schools: measuring organizational climate*: California: Sage Publications.

Hsieh, H., & Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505-528. doi:https://doi.org/10.1016/j.jsp.2011.07.001

Hughes, A. (2011). *The relationships between self-related perceptions, motivation, aspirations and achievements in an academic setting* (Doctoral dissertation). Staffordshire University, Stoke-on-Trent

Hugo, G. (1995). International labour migration and the family: some observations from Indonesia. *Asian and Pacific Migration Journal, 4*(2-3), 273-301.

Hui, C. (1988). Measurement of individualism-collectivism. *Journal of Research in Personality, 22*(1), 17-36. doi:https://doi.org/10.1016/0092-6566(88)90022-0

Indonesia, Republic of. (2003). *Undang-Undang No 20 Tahun 2003 tentang sistem pendidikan indonesia* Jakarta Retrieved from http://peraturan.go.id/common/dokumen/ln/2003/uu20-2003.pdf

Indonesia, Republic of. (2005). *Peraturan Pemerintah No 19 Tahun 2005 tentang Standar Nasional Pendidikan*. Jakarta Retrieved from https://peraturan.bpk.go.id/Home/Details/49369/pp-no-19-tahun-2005

Indonesia, Republic of. (2010). *Peraturan Pemerintah Republik Indonesia Nomor 17 Tahun 2010 tentang pengelolaan dan penyelenggaraan pendidikan* Jakarta Retrieved from http://peraturan.go.id/common/dokumen/ln/2010/pp17-2010bt.pdf

Iramaneerat, C., Smith Jr., E., & Smith, R. (2008). An introduction to rasch measurement. In J. Osborn (Ed.), *Best Practices in Quantitative Methods* (pp. 50-70). doi:10.4135/9781412995627

Irmawati. (2007). *Nilai-nilai yang mendasari motif-motif penentu keberhasilan suku batak toba.* (PhD). University of Indonesia, Jakarta.

Ishak, A., Osman, M., Mahaiyadin, M., Tumiran, M., & Anas, N. (2018). Examining unidimensionality of psychometric properties via Rasch model. *9*, 1462-1467.

Ivankova, N., Creswell, J., & Stick, S. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods, 18*(1), 3-20. doi:10.1177/1525822x05282260

Jalal, F., Samani, M., Chang, M., Stevenson, R., Ragatz, A., & Negara, S. (2009). *Teacher certification in Indonesia : a strategy for teacher quality improvement*. Indonesia: Ministry of National Education & World Bank

James, L. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*(2), 219-229. doi:10.1037/0021-9010.67.2.219

Jawas, U. (2008). Model kepemimpinan kepala sekolah dalam pelaksanaan manajemen berbasis sekolah di tingkat Sekolah Menengah Atas Negeri di kota Malang. *Malang: Universitas Muhammadiyah Malang. motivation and job satisfaction. Asian Journal of Technology Innovation, 19*(2), 233-247.

Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Book Inc.

Jennings, J., & Bearak, J. (2014). "Teaching to the Test" in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance. *Educational Researcher, 43*(8), 381-389. doi:10.3102/0013189X14554449

Jia, Y., Way, N., Ling, G., Yoshikawa, H., Chen, X., Hughes, D., Ke, X., & Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: a comparison of chinese and american adolescents. *Child Dev, 80*(5), 1514-1530. doi:10.1111/j.1467-8624.2009.01348.x

Jiang, Y., Song, J., Lee, M., & Bong, M. (2013). Self-efficacy and achievement goals as motivational links between perceived contexts and achievement. *Educational Psychology, 34*(1), 92-117. doi:10.1080/01443410.2013.863831

Johnson, R., & Onwuegbuzie, A. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher, 33*(7), 14-26. doi:10.3102/0013189x033007014

Joncas, M., & Foy, P. (2012). Sample Design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Jones, K. (2019). *Can I treat data of TIMSS study as three level. Retrieved from https://www.researchgate.net/post/Can-I-treat-data-of-TIMSS-study-as-three-level.*

Jones, R. (1995). *The child-school interface : environment and behaviour*. London: Cassell.

Kaluge, L. (1998). *Some factors related to educational attainment in Indonesian primary schools.* (PhD Thesis). University of London, London.

Kaluge, L., Setiasih, & Tjahjono, H. (2004). The quality improvement of Primary children learning through a school-based programme in Indonesia. In *Surabaya Indonesia*: Universitas Surabaya.

Karpicke, J., Blunt, J., & Smith, M. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology, 7*, 350.

Kholily, A. (2017). Dilemma of full day school in islamic education in indonesia. *Didaktika Religia, 5*(2), 345-360.

King, N. (1998). Template analysis. In G. Symon & C. Cassell (Eds.), *Qualitative methods and analysis in organizational research: A practical guide* (pp. 118-134). Thousand Oaks, CA: Sage Publications Ltd.

Kirkpatrick, R., & Zang, Y. (2011). The negative influences of exam-oriented education on Chinese high school students: Backwash from classroom to child. *Language testing in Asia, 1*(3), 36.

Knapik, M. (2006). The qualitative research interview: Participants' responsive participation in knowledge making. *International Journal of Qualitative Methods, 5*(3), 77-93.

Knight, J. (1985). *Building a Positive School Climate*. Paper presented at the the Annual Meeting of the Midwest Educational Research Association, Chicago.

Knuver, A., & Brandsma, H. (1993). Cognitive and Affective Outcomes in School Effectiveness Research. *School Effectiveness and School Improvement, 4*(3), 189-204. doi:10.1080/0924345930040302

Koklanaris, N., MacKenzie, A., Fino, M., Arslan, A., & Seubert, D. (2008). Debate preparation/participation: an active, effective learning tool. *Teaching and learning in medicine*, *20*(3), 235-238.

Konold, T., Cornell, D., Huang, F., Meyer, P., Lacey, A., Nekvasil, E., . . . Shukla, K. (2014). Multilevel Multi-Informant Structure of the Authoritative School Climate Survey. *School Psychology Quarterly, 29*(3), 238-255. doi:10.1037/spq0000062

Koplow, L. (2002). *Creating schools that heal: Real-life solutions*: Teachers College Press.

Korn, E., & Graubard, B. (1995). Analysis of Large Health Surveys: Accounting for the Sampling Design. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 158*(2), 263-295. doi:10.2307/2983292

Kreft, I., & Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.

Kreft, I., Leeuw, J., & Aiken, L. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*(1), 1-21.

Kusumawardhani, P. (2017). Does teacher certification program lead to better quality teachers? Evidence from Indonesia. *Education Economics, 25*(6), 590-618. doi:10.1080/09645292.2017.1329405

Kutsyuruba, B., Klinger, D., & Hussain, A. (2015). Relationships among school climate, school safety, and student achievement and well-being: a review of the literature. *Review of Education. 3*(2), 103-135. doi:10.1002/rev3.3043

Kwartolo, Y. (2002). Catatan kritis tentang kurikulum berbasis kompetensi. *Jurnal Pendidikan Penabur, 1*(1), 75-85.

Kyriakides, L. (2006). Using international comparative studies to develop the theoretical framework of educational effectiveness research: A secondary analysis of TIMSS 1999 data. *Educational Research and Evaluation*, *12*(6), 513-534.

Kyriakides, L., Campbell, R., & Gagatsis, A. (2000). The significance of the classroom effect in primary schools: An application of Creemers' comprehensive model of educational effectiveness. *School effectiveness and school improvement, 11*(4), 501-529.

Kyriakides, L., & Creemers, B. (2008). A longitudinal study on the stability over time of school and teacher effects on student outcomes. *Oxford Review of Education, 34*(5), 521-545.

Kyriakides, L., & Creemers, B. (2009). The effects of teacher factors on different outcomes: two studies testing the validity of the dynamic model. *Effective Education, 1*(1), 61-85. doi:10.1080/19415530903043680

Kyriakides, L., & Creemers, B. (2012). School policy on teaching and school learning environment: direct and indirect effects upon student outcome measures. *Educational Research and Evaluation, 18*(5), 403-424. doi:10.1080/13803611.2012.689716

Kyriakides, L., Creemers, B., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education, 25*(1), 12-23. doi:10.1016/j.tate.2008.06.001

Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: implications for theory and research. *British Educational Research Journal, 36*(5), 807-830. doi:10.1080/01411920903165603

Kyriakides, L., Creemers, B., Muijs, D., Rekers-Mombarg, L., Papastylianou, D., Van Petegem, P., & Pearson, D. (2013). Using the dynamic model of educational effectiveness to design strategies and actions to face bullying. *School Effectiveness and School Improvement, 25*(1), 83-104. doi:10.1080/09243453.2013.771686

Lai, S., Ye, R., & Chang, K. (2008). Bullying in middle schools: An Asian-Pacific regional study. *Asia Pacific Education Review, 9*(4), 503-515.

Larson, L., & Lovelace, M. (2013). Evaluating the efficacy of questioning strategies in lecture-based classroom environments: Are we asking the right questions?. *Journal on Excellence in College Teaching, 24*(1).

Lauder, H., Jamieson, I., & Wikeley, H. (1998). Model of effective schools: limits and capabilities. In R. Slee, S. Tomlinson, & G. Weiner (Eds.), *School Effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements*. London: Taylor & Francis e-Library.

Laukaityte, I., & Wiberg, M. (2017). Using plausible values in secondary analysis in large-scale assessments. *Communications in Statistics - Theory and Methods, 46*(22), 11341-11357. doi:10.1080/03610926.2016.1267764

Laukaityte, I., & Wiberg, M. (2018). Importance of sampling weights in multilevel modeling of international large-scale assessment data. *Communications in Statistics - Theory and Methods, 47*(20), 4991-5012. doi:10.1080/03610926.2017.1383429

Leeuw, J., & Meijer, E. (2008). Introduction to Multilevel Analysis. In J. Leeuw & E. Meijer (Eds.), *Handbook of Multilevel Analysis* (pp. 1-75): New York: Springer.

Legge, J., McDivitt, J., Leinbach, T., Mohamad, G., Wolters, O., & Adam, A. (2016). *Indonesia*. Retrieved from https://www.britannica.com/place/Indonesia

Lenkeit, J. (2013). Effectiveness measures for cross-sectional studies: a comparison of value-added models and contextualised attainment models. *School Effectiveness and School Improvement, 24*(1), 39-U38. doi:10.1080/09243453.2012.680892

Lenkeit, J., & Caro, D. (2014). Performance status and change – measuring education system effectiveness with data from PISA 2000–2009. *Educational Research and Evaluation, 20*(2), 146-174. doi:10.1080/13803611.2014.891462

Lepola, J. (2000). *Motivation in early school years : development patterns and cognitive consequences*. Turku: Turun Yliopisto.

Lewis, R. (1997). *Learning Styles in Transition: A Study of Indonesian Students*. Paper presented at the the Annual Meeting of the Japan Association of Language Teachers, Hamamatsu, Japan.

Liem, G. (2016). The effects of culture and sex on students' approaches to learning: inspiring insights from David Watkins' intellectual inquiries. In R. B. King & A. B. I. Bernanrdo (Eds.), *The psychology of Asian learners* (pp. 217-233). Boston: Springer.

Liem, G., Martin, A., Nair, E., Bernardo, A., & Prasetya, P. (2009). Cultural factors relevant to secondary school students in Australia, Singapore, the Philippines and Indonesia: Relative differences and congruencies. *Journal of Psychologists and Counsellors in Schools, 19*(2), 161-178.

Lin, B., & Crawley, F. (1987). Classroom climate and science-related attitudes of junior-high-school students in taiwan. *Journal of Research in Science Teaching, 24*(6), 579-591. doi:10.1002/tea.3660240607

Linacre, J. (2011a). A user's guide to Winsteps-Ministep: rasch-model computer programs. Program manual 3.73.0. In. Beaverton, Oregon: Winsteps.com.

Linacre, J. (2011b). Winsteps Rasch Measurement. Beaverton, Oregon: Winsteps.com.

Lincoln, Y., Lynham, S., & Guba, E. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (Vol. 4, pp. 97-128). London: Sage Publication.

Lipscomb, M. (2011). Challenging the coherence of social justice as a shared nursing value. *Nursing Philosophy, 12*(1), 4-11.

Little, R., & Rubin, D. (2019). *Statistical analysis with missing data*. 3rd edition. Chichester: John Wiley & Sons.

Liu, C. (2004). *The relationship between school culture and student achievement in Arizona elementary public schools.* (Doctoral). The University of Arizona.

Louis, K. (2006). Changing the Culture of Schools: Professional Community, Organizational Learning, and Trust. *Journal of School Leadership, 16*(5), 477-489. doi:10.1177/105268460601600502

Loukas, A. (2007). What is school climate. *Leadership compass, 5*(1), 1-3.

Loukas, A., & Murphy, J. (2007). Middle school student perceptions of school climate: Examining protective functions on subsequent adjustment problems. *Journal of School Psychology, 45*(3), 293-309. doi:https://doi.org/10.1016/j.jsp.2006.10.001

Luyten, H., Visscher, A., & Witziers, B. (2005). School effectiveness research: From a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement, 16*(3), 249-279. doi:10.1080/09243450500114884

Ma, X. (2001). Stability of School Academic Performance across Subject Areas. *Journal of Educational Measurement, 38*(1), 1-18. Retrieved from http://www.jstor.org/stable/1435436

Ma'zumi, M., Jakaria, J., & Research, T. (2012). Contribution of Madrasah to the Development of the Nation Character. *International Journal of Scientific, 1*(11), 37-39. doi: 10.1.1.300.2601

Mackatiani, C. (2017). Influence of Examinations Oriented Approaches on Quality Education in Primary Schools in Kenya. *Journal of Education and Practice, 8*(14), 51-58.

Magnis-Suseno, F. (1997). *Javanese Ethics and World-View. The Javanese Idea of the Good Life*. Jakarta: Penerbit PT Gramedia Pustaka Utama.

Makwinja-Morara, V. (2009). Female dropouts in Botswana junior secondary schools. *Educational studies, 45*(5), 440-462.

Marambe, K., Vermunt, J., & Boshuizen, H. (2012). A cross-cultural comparison of student learning patterns in higher education. *Higher Education, 64*(3), 299-316.

Margianti, E., Fraser, B., & Aldridge, J. (2001). *Classroom environment and students' outcomes among university computing students in Indonesia.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Marks, H. (2000). Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research Journal, 37*(1), 153-184. doi:10.2307/1163475

Markus, K. (2008). Constructs, Concepts and the Worlds of Possibility: Connecting the Measurement, Manipulation, and Meaning of Variables. *Measurement: Interdisciplinary Research and Perspectives, 6*(1-2), 54-77. doi:10.1080/15366360802035513

Marsh, H. (1990a). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panle analysis. *Journal of Educational Psychology, 82*(4), 646-656.

Marsh, H. (1990b). A Multidimensional, Hierarchical Model of Self-Concept: Theoretical and Empirical Justification. *Educational Psychology Review, 2*(2), 77-172. doi:Doi 10.1007/Bf01322177

Marsh, H. , Ludtke, O., Nagengast, B., Trautwein, U., Morin, A., Abduljabbar, A., & Koller, O. (2012). Classroom Climate and Contextual Effects: Conceptual and Methodological Issues in the Evaluation of Group-Level Effects. *Educational Psychologist, 47*(2), 106-124. doi:10.1080/00461520.2012.670488

Marsh, H., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self concept. *Personality and social psychology bulletin, 34*(4), 542-552.

Marsh, H., & Martin, A. (2011). Academic self-concept and academic achievement: relations and causal ordering. *Br J Educ Psychol, 81*(Pt 1), 59-77. doi:10.1348/000709910X503501

Marsh, H., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107-123.

Marshall, B. (2017). The politics of testing. *English in Education, 51*(1), 27-43. doi:10.1111/eie.12110

Martin, M., & Mullis, I. (Eds.). (2013). *Methods and Procedures in TIMSS and PIRLS 2011*. Boston, USA: TIMSS & PIRLS International Study Center & International Association for the Evaluation of Educational Achievement (IEA).

Martin, M., Mullis, I., Foy, P., & Arora, A. (2012). Creating and Interpreting the TIMSS and PIRLS 2011 Context Questionnaire Scales. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Marzano, R., Marzano, J., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher*. ASCD.

Mason, E. (1973). Teachers' observations and expectations of boys and girls as influenced by biased psychological reports and knowledge of the effects of bias. *Journal of Educational Psychology, 65*(2), 238-243. doi:10.1037/h0034979

Maulana, R., Helms-Lorenz, M., Irnidayanti, Y., & van de Grift, W. (2016). Autonomous motivation in the Indonesian classroom: Relationship with teacher support through the lens of self-determination theory. *25*(3), 441-451.

Maulana, R., Opdenakker, M., den Brok, P., & Bosker, R. (2011). Teacher–student interpersonal relationships in Indonesia: profiles and importance to student

motivation. *Asia Pacific Journal of Education, 31*(1), 33-49. doi:10.1080/02188791.2011.544061

Maxwell, J., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. *Handbook of mixed methods in social & behavioral research*, 145-168.

McSweeney, B. (2002). Hofstede's Model of National Cultural Differences and their Consequences: A Triumph of Faith - a Failure of Analysis. *Human Relations, 55*(1), 89-118. doi:10.1177/0018726702551004

Meeus, W., Oosterwegel, A., & Vollebergh, W. (2002). Parental and peer attachment and identity development in adolescence. *J Adolesc, 25*(1), 93-106. doi:10.1006/jado.2001.0451

Meraviglia, M., Becker, H., Rosenbluth, B., Sanchez, E., & Robertson, T. (2003). The Expect Respect Project: Creating a positive elementary school climate. *Journal of Interpersonal Violence, 18*(11), 1347-1360.

Merriam, S. (1988). *Case study research in education : a qualitative approach*. London: Jossey-Bass.

Merriam, S. (2002). *Qualitative research in practice : examples for study and discussion*. San Francisco: Jossey-Bass.

Miles, M. (1965). Planned change and organizational health: Figure and ground In R. O. Carlson, A. J. Gallaher, M. b. Miles, R. J. Pellegrin, & E. M. Rogers (Eds.), *change processes in the public schools*. Oregon: The Center For The Advanced Study of Educational Administration.

Miles, M., Huberman, A., & Saldaña, J. (2014). *Qualitative data analysis : a methods sourcebook* (3rd ed.). Thousand Oaks, Calif. ; London: Sage Publications.

Miller, S., & Fredericks, J. (1990). The False Ontology of School Climate Effects. *Educational Theory, 40*(3), 333-342. doi:10.1111/j.1741-5446.1990.00333.x

Ministry of Education and Culture, Republic of Indonesia. (2005). *Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 29 Tahun 2005 tentang Badan Akreditasi Nasional Sekolah/Madrasah*. Jakarta: Ministry of Education and Culture

Ministry of Education and Culture, Republic of Indonesia. (2012). *Indonesia Educational Statistics In Brief/ Ringkasan Statistik Pendidikan Indonesia 2011/2012*. Jakarta: Ministry of Education and Culture Retrieved from http://kemdikbud.go.id/kemdikbud/dokumen/BukuRingkasanDataPendidikan/Final-In-Brief-1112.pdf

Ministry of Education and Culture, Republic of Indonesia. (2013a). *Overview of the Education Sector in Indonesia 2012: Achievements and Challenges.* Jakarta: Ministry of Education and Culture.

Ministry of Education and Culture, Republic of Indonesia. (2013b). *Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 67 Tahun 2013 tentang kerangka dasar dan struktur kurikulum sekolah dasar/madrasah ibtidaiyah Jakarta* Retrieved from http://simpuh.kemenag.go.id/regulasi/permendikbud_67_13_lampiran.pdf

Ministry of Education and Culture, Republic of Indonesia. (2017a). Indonesia Educational Statistic in Brief 2016/2017. In *Yearly*. Jakarta: Center for educational data and statistics and culture.

Ministry of Education and Culture, Republic of Indonesia. (2017b). *Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2017 tentang penilaian hasil belajar oleh pemerintah dan penilaian hasil belajar oleh satuan pendidikan*. Jakarta Retrieved from http://simpuh.kemenag.go.id/regulasi/permendikbud_03_17.pdf

Ministry of Education and Culture, Republic of Indonesia. (2018). *Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 35 Tahun 2018 tentang perubahan atas peraturan menteri pendidikan dan kebudayaan nomor 58 tahun 2014 tentang kurikulum 2013 sekolah menengah pertama/madrasah tsanawiyah*. Jakarta Retrieved from http://peraturan.go.id/common/dokumen/bn/2018/bn1690-2018.pdf

Ministry of Religious Affairs, Republic of Indonesia. (2014a). *Keputusan Menteri Agama Republik Indonesia Nomor 207 Tahun 2014 tentang kurikulum madrasah*. Jakarta Retrieved from http://simpuh.kemenag.go.id/regulasi/kma_207_14.pdf

Ministry of Religious Affairs, Republic of Indonesia. (2014b). *Madrasah lebih baik (Madrasah: the better choice of education)*. Jakarta: Direktorat Pendidikan Madrasah Retrieved from http://madrasah.kemenag.go.id/files/PROFIL_MADRASAH_LENGKAP.pdf

Mo, Y., Singh, K., & Chang, M. (2013). Opportunity to learn and student engagement: A HLM study on eighth grade science achievement. *Educational Research for Policy and Practice*, *12*(1), 3-19.

Mohammadpour, E. (2013). A three-level multilevel analysis of Singaporean eighth-graders science achievement. *Learning and Individual Differences, 26*(0), 212-220. doi:10.1016/j.lindif.2012.12.005

Mooij, M. (2014). International and Cross-Cultural Consumer Behavior. In H. Cheng (Ed.), *The Handbook of International Advertising Research* (pp. 125-148).

Moos, R. (1980). Evaluating classroom learning environments. *Studies in Educational Evaluation, 6*, 239-252.

Moos, D., & Ringdal, A. (2012). Self-regulated learning in the classroom: A literature review on the teacher's role. *Education Research International, 2012*.

Moreira, B., Pinto, T., Starling, D., & Jaeger, A. (2019, February). Retrieval practice in classroom settings: a review of applied research. In *Frontiers in Education* (Vol. 4, p. 5). Frontiers.

Morgan, D. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research, 1*(1), 48-76. doi:10.1177/2345678906292462

Morgan, D. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry, 20*(8), 1045-1053. doi:10.1177/1077800413513733

Morin, A., Marsh, H., Nagengast, B., & Scalas, L. (2013). Doubly Latent Multilevel Analyses of Classroom Climate: An Illustration. *The Journal of Experimental Education, 82*(2), 143-167. doi:10.1080/00220973.2013.769412

Morley, L., & Rassool, N. (2002). *School effectiveness: Fracturing the discourse*, New York: Routledge.

Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nurs Res, 40*(2), 120-123.

Mortimore, P. (1988). *School matters : the junior years*. Wells: Open Books.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1989). A study of effective junior schools. *International Journal of Educational Research, 13*(7), 753-768. doi:https://doi.org/10.1016/0883-0355(89)90026-8

Moustakas, C. (1994). *Phenomenological research methods*, Thousand Oaks: Sage.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2010). Improving Schools in Socioeconomically Disadvantaged Areas – A Review of Research Evidence. doi:School Effectiveness and School Improvement, June 01, 2004, Vol. 15, No. 2, pp. 149–175

Muijs, D., & Reynolds, D. (2003). Student Background and Teacher Effects on Achievement and Attainment in Mathematics: A Longitudinal Study. *Educational Research and Evaluation, 9*(3), 289-314. doi:10.1076/edre.9.3.289.15571

Mullis, I., Drucker, K., Preuschoff, C., Arora, A., & Stanco, G. (2012). Assessment Framework and Instrument Development. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I., Martin, M., Foy, P., & Arora, A. (2012). *TIMSS 2011 International result in mathematics*. Boston: TIMSS & PIRLS International Study Center.

Mullis, I., Martin, M., Robitaille, D., & Foy, P. (2009). *TIMSS advanced 2008 international report: Findings from IEA's study of achievement in advanced mathematics and physics in the final year of secondary school*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Muñoz-Chereau, B. (2013). *Searching for fairer ways of comparing Chilean secondary schools performance: a mixed methods study investigating contextual value added approaches.* (PhD). University of Bristol, Bristol. Retrieved from https://doi.org/10.1080/09243453.2018.1503604

Muñoz-Chereau, B. (2019). Exploring gender gap and school differential effects in mathematics in Chilean primary schools. *School Effectiveness and School Improvement, 30*(2), 83-103. doi:10.1080/09243453.2018.1503604

Mustafa, A. (2014). *Upaya Meningkatkan Kemampuan Berpikir Kritis dan Kreatif serta Self-Efficacy dalam Pembelajaran Matematika Melalui Discovery Learning.* (Doctoral Dissertation). Universitas Pendidikan Indonesia,

Muttaqin, T., Wittek, R., Heyse, L., & van Duijn, M. (2019). The achievement gap in Indonesia? Organizational and ideological differences between private Islamic

schools. *School Effectiveness and School Improvement*, 1-31. doi:10.1080/09243453.2019.1644352

Myers, K., & Goldstein, H. (1998). Who's failing. *No Quick Fixes: perspectives on schools in difficulty*, 175-188.

Napitupulu, E. (2012). Ujian Nasional matikan kreativitas. *Kompas*. Retrieved from https://internasional.kompas.com/read/2012/07/07/12102361/ujian.nasional.matikan.kreativitas

Newhouse, D., & Beegle, K. (2006). The effect of school type on academic achievement evidence from indonesia. *Journal of Human Resources, 41*(3), 529-557.

Noesjirwan, J. (1978). A rule-based analysis of cultural differences in social behaviour: Indonesia and Australia. *International Journal of Psychology, 13*(4), 305-316.

Nor, M., & Malim, M. (2014). Revisiting Islamic education: the case of Indonesia. *Journal for Multicultural Education, 8*(4), 261-276.

Novera, I. (2004). Indonesian postgraduate students studying in Australia: An examination of their academic, social and cultural experiences. *International Education Journal*, *5*(4), 475-487.

O'Mara, A., Marsh, H., Craven, R., & Debus, R. (2006). Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis. *Educational Psychologist, 41*(3), 181-206. doi:10.1207/s15326985ep4103_4

Oberheim, E. (2016). Rediscovering Einstein's legacy: How Einstein anticipates Kuhn and Feyerabend on the nature of science. *Studies in History and Philosophy of Science Part A, 57*, 17-26. doi:https://doi.org/10.1016/j.shpsa.2015.11.005

OECD. (2009). *PISA 2009 assessment framework : key competencies in reading, mathematics and science*. Paris: OECD.

OECD. (2013). *PISA 2012 Results: Ready to Learn (Volume III)*. Paris: OECD.

OECD & ADB. (2015). *Reviews of National Policies for Education: Education in Indonesia - Rising to the Challenge*. Retrieved from: https://think-asia.org/handle/11540/3308

Oerter, R., Oerter, R., Agostiani, H., Kim, H. O., & Wibowo, S. (1996). The concept of human nature in East Asia: Etic and emic characteristics. *Culture & Psychology*, *2*(1), 9-51.

Opdenakker, M., & Damme, J. (2005). Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools. *School Effectiveness and School Improvement*, 17 (1), 87–117

Opdenakker, M., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement, 11*(2), 165-196. doi:Doi 10.1076/0924-3453(200006)11:2;1-Q;Ft165

Opdenakker, M., Van Damme, J., De Fraine, B., Van Landeghem, G., & Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *School Effectiveness and School Improvement, 13*(4), 399-427. doi:DOI 10.1076/sesi.13.4.399.10283

Oros, A. (2007). Let's debate: Active learning encourages student participation and critical thinking. *Journal of Political Science Education*, *3*(3), 293-311.

Othman, M., & Muijs, D. (2013). Educational quality differences in a middle-income country: the urban-rural gap in Malaysian primary schools. *School Effectiveness and School Improvement, 24*(1), 1-18. doi:10.1080/09243453.2012.691425

Pajares, F., & Urdan, T. (2002). *Academic motivation of adolescents*. Greenwich: Information Age Pub.

Panayides, P., Robinson, C., & Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal, 36*(4), 611-626.

Panayiotou, A., Kyriakides, L., & Creemers, B. (2016). Testing the validity of the dynamic model at school level: A European study. *School Leadership & Management, 36*(1), 1-20.

Park, J., & Niyozov, S. (2008). Madrasa education in South Asia and Southeast Asia: current issues and debates. *Asia Pacific Journal of Education, 28*(4), 323-351. doi:Pii 90663514010.1080/02188790802475372

Parker, L., & Raihani, R. (2009). Governing madrasah. *Policy Briefs*, 1-10.

Parker, P., Marsh, H., Ciarrochi, J., Marshall, S., & Abduljabbar, A. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*(1), 29-48. doi:10.1080/01443410.2013.797339

Paschal, B. (1968). The role of self-concept in achievement. *The Journal of Negro Education, 37*(4), 392-396.

Pashler, H., Rohrer, D., Cepeda, N., & Carpenter, S. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic bulletin & review, 14*(2), 187-193.

Patrick, H., Ryan, A., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99*(1), 83-98. doi:10.1037/0022-0663.99.1.83

Payne, A. (2018). Creating and Sustaining a Positive and Communal School Climate: Contemporary Research, Present Obstacles, and Future Directions. National Institute of Justice Report. *National Institute of Justice*.

Payne, G. & Payne, J. (2004). Longitudinal and cross-sectional studies. In *Key concepts in social research* (pp. 144-148). London: Sage Publications, Ltd, https://www.doi.org/10.4135/9781849209397

Pepinsky, T. (2018). A Note on Listwise Deletion versus Multiple Imputation. *Political Analysis, 26*(4), 480-488. doi:10.1017/pan.2018.18

Perry, K., & Weinstein, R. (1998). The social context of early schooling and children's school adjustment. *Educational Psychologist, 33*(4), 177-194.

Peterson, K., & Deal, T. (2009). *The shaping school culture fieldbook* (2 ed.). San Francisco, CA: Jossey-Bass Inc.

Phelps, R. (2011). Teach to the Test? *The Wilson Quarterly (1976-), 35*(4), 38-42.

Pollitt, A., & Ahmed, A. (2001). *Science or Reading?: How students think when answering TIMSS questions. Retrieved February*. Paper presented at the International Association for Educational Assessment, Rio de Janeiro.

Prosser, J. (1999). The Evolution of school culture research. In J. Prosser (Ed.), *School culture* (pp. xvii, 182 p.). London: Paul Chapman.

Rabasa, A. (2005). Islamic Education in Southeast Asia. *Current trends in Islamist ideology, 2*, 97-108.

Rachmadiana, M. (2004). Mencium tangan, membungkukkan badan: Etos Budaya Sunda, Yogyakarta, Madura. *Humanitas: Jurnal Psikologi Indonesia, 1*(2).

Ragatz, A., Iskandar, S., Kesuma, R., Sugiarti, S., & Ragatz, A. B. (2015). *Indonesia - A video study of teaching practices in TIMSS eighth grade mathematics classrooms : understanding what teaching practices are used, why they are used and how they relate to student learning*. Main Report. Jakarta: The World Bank

Rahmi, D., Wibisono, B., & Setiawan, B. (2001). Rukun and Gotong Royong: Managing Public Places in an Indonesian Kampung. In P. Miao (Ed.), *Public Places in Asia Pacific Cities.* (Vol. 60). Dordrecht: Springer.

Rasbash, J., Steele, F., Browne, W., & Goldstein, H. (2012). *A user's guide to MLwiN : version 2.26* (Updated [ed.]. ed.). Bristol: Centre for Multilevel Modelling, University of Bristol.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models : applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage.

Raudenbush, S., & Willms, J. (1995). The estimation of school effects. *Journal of educational behavioral statistics, 20*(4), 307-335.

Reaves, S., McMahon, S., Duffy, S., & Ruiz, L. (2018). The test of time: A meta-analytic review of the relation between school climate and problem behavior. *Aggression and Violent Behavior, 39*, 100-108. doi:https://doi.org/10.1016/j.avb.2018.01.006

Reezigt, G., Guldemond, H., & Creemers, B. (1999). Empirical validity for a comprehensive model on educational effectiveness. *School effectiveness and school improvement, 10*(2), 193-216.

Reid, K., Hopkins, D., & Holly, P. (1987). *Towards the effective school.* Michigan: Basil Blackwell.

Renne, C. (1996). Structuring classroom lessons: Attempts to incorporate student questions and initiatives during math lessons. *Teacher Education Quarterly*, 5-18.

Revina, S. (2017). *Influence of culture on the adaptation of realistic mathematics education in indonesia.* (PhD Doctoral). University of Hongkong, Hongkong.

Reyes, M., Brackett, M., Rivers, S., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700-712. doi:10.1037/a0027268

Reynolds, D. (1982). The Search for Effective Schools. *School Organisation, 2*(3), 215-237. doi:10.1080/0260136820020302

Reynolds, D., Bollendorf, R., Creemers, B., & Hopkins, D. (1996). *Making good schools: Linking school effectiveness and school improvemen.t.* Hove: Psychology Press.

Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement, 25*(2), 197-230. doi:10.1080/09243453.2014.885450

Reynolds, D., & Teddlie, C. (2002). The processes of school effectiveness. In D. Reynolds & C. Teddlie (Eds.), *The international handbook of school effectiveness research* (pp. 148-173). New York: Routledge.

Reynolds, D., Teddlie, C., Chapman, C., & Stringfield, S. (2015). Effective school processes. In *The Routledge International Handbook of Educational Effectiveness and Improvement*. New York: Routledge.

Riddell, A. (1997). Assessing Designs for School Effectiveness Research and School Improvement in Developing Countries. *Comparative Education Review, 41*(2), 178-204. Retrieved from http://www.jstor.org/stable/1188835

Riessman, C. (1993). *Narrative analysis* (Vol. 30). Thousand Oaks: Sage.

Rimm-Kaufman, S., & Sandilos, L. (2011). *Improving students' relationships with teachers to provide essential supports for learning.* Retrieved from: https://www.apa.org/education/k12/relationships

Ringle, C., Wende, S., & Becker, J. (2015). SmartPLS 3.

Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. Burgess (Eds.), *Analyzing qualitative data*. Canada: Routledge.

Robson, C., & McCartan, K. (2016). *Real world research*. Chichester: John Wiley & Sons.

Roeser, R., Eccles, J., & Sameroff, A. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *The Elementary School Journal, 100*(5), 443-471.

Roeser, R., Midgley, C., & Urdan, T. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*(3), 408.

Rosenkvist, M. (2010). Using Student Test Results for Accountability and Improvement: A Literature Review. OECD Education Working Papers, No. 54. *OECD Publishing (NJ1)*.

Rosser, A. (2018). *Beyond access: making Indonesia's education system work.* Lowy Institute

Rosyad, R. (2007). *A quest for true Islam: A Study of the Islamic resurgence movement among the youth in Bandung, Indonesia* (p. 110). ANU Press.

Rowe, K., & Hill, P. (1998). Modeling Educational Effectiveness in Classrooms: The Use of Multi-Level Structural Equations to Model Students' Progress. *Educational Research and Evaluation, 4*(4), 307-347. doi:10.1076/edre.4.4.307.6953

Rubie-Davies, C. (2007). Classroom interactions: Exploring the practices of high- and low-expectation teachers. *77*(2), 289-306. doi:10.1348/000709906x101601

Rubie-Davies, C. (2014). *Becoming a high expectation teacher: Raising the bar*. London: Routledge.

Rubin, L., Witkiewitz, K., Andre, J., & Reilly, S. (2007). Methods for Handling Missing Data in the Behavioral Neurosciences: Don't Throw the Baby Rat out with the Bath Water. *J Undergrad Neurosci Educ, 5*(2), A71-77.

Rudasill, K., Snyder, K., Levinson, H., & Adelson, J. (2018). Systems View of School Climate: a Theoretical Framework for Research. *Educational Psychology Review, 30*(1), 35-60. doi:10.1007/s10648-017-9401-y

Rudhito, M., & Prasetyo, D. (2016). Pengembangan Soal Matematika Model TIMSS untuk Mendukung Pembelajaran Matematika SMP Kelas VII Kurikulum 2013. *Jurnal Cakrawala Pendidikan, 35*(1).

Rust, K. (2014). Sampling, weighting, and variance estimation in lage-scale assessment. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. Boca Raton: Taylor & Francis Group.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher, 39*(2), 142-151.

Rutkowski, L., & Rutkowski, D. (2009). Trends in TIMSS responses over time: evidence of global forces in education? *Educational Research and Evaluation, 15*(2), 137-152. doi:10.1080/13803610902784352

Rutter, M. (1979). *Fifteen thousand hours : secondary schools and their effects on children*. London: Open Books.

Rutter, M., & Maughan, B. (2002). School Effectiveness Findings 1979–2002. *Journal of School Psychology, 40*(6), 451-475. doi:10.1016/s0022-4405(02)00124-3

Ryle, G. (2015). IX.—Categories. *Proceedings of the Aristotelian Society, 38*(1), 189-206. doi:10.1093/aristotelian/38.1.189 %J Proceedings of the Aristotelian Society

Salim, M. (2011). *Exploring issues of school effectiveness and self-evaluation at the system and school levels in the context of Zanzibar*. (PhD thesis). University of Bristol, Bristol, UK.

Sammons, P. (1999). *School effectiveness : coming of age in the twenty-first century*. Lisse: Swets & Zeitlinger.

Sammons, P. (2009). The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools. *School Effectiveness and School Improvement, 20*(1), 123-129. doi:10.1080/09243450802664321

Sammons, P. (2010). The Contribution of Mixed Methods to Recent Research on Educational Effectiveness. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods Research* (pp. 697-723). London, UK: Sage Publication.

Sammons, P., Nuttall, D., & Cuttance, P. (1993). Differential School Effectiveness: results from a reanalysis of the Inner London Education Authority's Junior School Project Data. *British Educational Research Journal, 19*(4), 381-405. doi:10.1080/0141192930190407

Sammons, P., Nuttall. D., Cuttance, P., & Thomas, S. (1995). Continuity of school effects: A longitudinal analysis of primary and secondary school effects on GSCE performance. *School effectiveness and school improvement,* 6(4), 285-307.

Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging links : effective schools and effective departments*. London: Paul Chapman.

Sammons, P., Thomas, S., Mortimore, P., Owen, C., & Pennell, H. (1994). *Assessing school effectiveness : Developing measures to put school performance in context*. London: Office for Standards in Education [OFSTED].

Sammons, P., Thomas, S., Mortimore, P., Walker, A., Cairns, R., & Bausor, J. (1998). Understanding Differences in Academic Effectiveness: Practitioners' Views∗. *School Effectiveness and School Improvement, 9*(3), 286-309. doi:10.1080/0924345980090302

Samosir, M. (2008). *The Effects of Decentralization on Education in Indonesia: Education for All?*. Master Theses in Public Policy and Human Development. Universiteit Maastrich. Maastrich Graduate School of Governance. The Netherlands.

Samuels, S. (1977). *Enhancing self-concept in early childhood : theory and practice*. New York: Human Sciences Press

Sanders, W., Wright, S., & Horn, S.(1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67. doi:10.1023/a:1007999204543

Sandoval-Hernandez, A. (2008). School effectiveness research: a review of criticisms and some proposals to address them. *Educate~, 1*(1), 31.

Saputro, B., Suryadi, D., Rosjanuardi, R., & Kartasasmita, B. (2018). Analysis of students' errors in responding to TIMSS domain algebra problem. *Journal of Physics: Conference Series, 1088*, 012031. doi:10.1088/1742-6596/1088/1/012031

Saxebøl, T. (2002). *The Madurese ulama as patrons: a case study of power relations in an Indonesian community.* (Master thesis). University of Oslo, Norwegia.

Schafer, J., & Graham, J. (2002). Missing data: our view of the state of the art.*Psychological Methods. 7*(2), 147.

Scheerens, J. (1990). School Effectiveness Research and the Development of Process Indicators of School Functioning. *School Effectiveness and School Improvement, 1*(1), 61-80. doi:10.1080/0924345900010106

Scheerens, J. (1992). *Effective schooling : research, theory and practice*. London: Cassell.

Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement, 24*(1), 1-38. doi:10.1080/09243453.2012.691100

Scheerens, J. (2014). School, teaching, and system effectiveness: some comments on three state-of-the-art reviews. *School Effectiveness and School Improvement, 25*(2), 282-290. doi:10.1080/09243453.2014.885453

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness.* Oxford: Pergamon Press.

Scheerens, J., & Bosker, R. J. (1996). *The foundations of educational effectiveness*. Oxford ; New York: Pergamon Elsevier.

Scheerens, J., Glas, C., & Thomas, S. (2003). *Educational evaluation assessment and monitoring : a systemic approach*. Lisse: Swets & Zeitlinger.

Scheerens, J., Hendriks, M., Luyten, H., Sleegers, P., & Cees, G. (2013). Productive time in education. *A Review of the Effectiveness of Teaching Time at School, Homework and Extended Time Outside School Hours. Enschede: University of Twente. Available online at: http://doc. utwente. nl/86371/(Accessed January 25, 2016).*

Schoen, L., & Teddlie, C. (2008). A new model of school culture: a response to a call for conceptual clarity. *School Effectiveness and School Improvement, 19*(2), 129-153. doi:10.1080/09243450802095278

Schunk, D. (1991). Self-Efficacy and Academic Motivation. *Educational Psychologist, 26*(3-4), 207-231. doi:10.1080/00461520.1991.9653133

Schunk, D. (2001). Self-efficacy: Educational aspects. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 13820-13822). Oxford, England: Pergamon.

Seaton, M., Parker, P., Marsh, H., Craven, R., & Yeung, A. (2014). The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology, 34*(1), 49-72. doi:10.1080/01443410.2013.825232

Seidman, I. (2006). *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences.* New York: Teachers College Press.

Shannon-Baker, P. (2015). Making Paradigms Meaningful in Mixed Methods Research. *Journal of Mixed Methods Research*. doi:10.1177/1558689815575861

Shavelson, R., Hubner, J., & Stanton, G. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research, 46*(3), 407-441. doi:10.3102/00346543046003407

Shernoff, D. (2013). *Optimal learning environments to promote student engagement.* Boston: Springer.

Shoemaker, P., Tankard Jr., J., & Lasorsa, D. (2004). How to Build Social Science Theories. In. doi:10.4135/9781412990110

Shultz, K., Hoffman, C., & Reiter-Palmon, R. (2005). Using Archival Data for I-O research: advantages, pitfalls, sources, and examples. *The Industrial Psychologist.42*(3). 31-37.

Signorini, P., Wiesemes, R., & Murphy, R. (2009). Developing alternative frameworks for exploring intercultural learning: a critique of Hofstede's cultural difference model. *Teaching in Higher Education, 14*(3), 253-264. doi:10.1080/13562510902898825

Sikkink, D. (2012). Religious School Differences in School Climate and Academic Mission: A Descriptive Overview of School Organization and Student Outcomes. *Journal of School Choice, 6*(1), 20-39. doi:10.1080/15582159.2012.651394

Silverman, D. (2013). *Doing Qualitative Research: A Practical Handbook.* London: Sage Publications

Simons, H. (2013). *Case study research in practice*. London: SAGE publications.

Slater, R., & Teddlie, C. (1992). Toward a Theory of School Effectiveness and Leadership. *School Effectiveness and School Improvement, 3*(4), 247-257. doi:10.1080/0924345920030402

Slavin, R., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research, 78*(3), 427-515.Doi 10.3102/0034654308317473

Smith, D., & Tomlinson, S. (1990). *The school effect: A study of multi racial comprehensives.* London: Policy Studies Institute

Smith, J., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis: Theory, method and research.* London :Sage publications.

Smith, J. (1983). Quantitative Versus Qualitative Research: An Attempt to Clarify the Issue. *Educational Researcher, 12*(3), 6-13. doi:10.3102/0013189X012003006

Smithson, J. (2000). Using and analysing focus groups: limitations and possibilities. *International journal of social research methodology, 3*(2), 103-119.

Sofo, F., Fitzgerald, R., & Jawas, U. (2012). Instructional leadership in Indonesian school reform: overcoming the problems to move forward. *School Leadership & Management, 32*(5), 503-522. doi:10.1080/13632434.2012.723616

Sopantini. (2014). *Reforming teaching practices in Indonesia: A case study of the implementation of active learning in primary schools in North Maluku.* (Ed.D Doctoral). University of Tasmania

Stake, R. (2006). *Multiple case study analysis*. New York: Guilford Press.

Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). Highlights from TIMSS and TIMSS Advanced 2015: Mathematics and Science Achievement of US Students in Grades 4 and 8 and in Advanced Courses at the End of High School in an International Context. NCES 2017-002. *National Center for Education Statistics*.

Steer, L., & Smith, K. (2015). Financing Education: Opportunities for Global Action. *Center for Universal Education at The Brookings Institution*.

Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement, 21*(3), 289-314. doi:10.1080/09243451003732651

Strand, S. (2016). Do some schools narrow the gap? Differential school effectiveness revisited. *Review of Education, 4*(2), 107-144. doi:10.1002/rev3.3054

Stringfield, S., & Teddlie, C. (2011). School effectiveness research, 1932–2008, including a call for future research. In *International handbook of teache* (pp. 379-388).

Stevenson, A. (ed.), 2010, *Oxford Dictionary of English*, third edition, Oxford: Oxford University Press.

Sumintono, B. (2006). *Decentralized centralism: School based management policies and practices at state secondary schools in Mataram, Lombok, Indonesia* (PhD Thesis). Victoria University of Wellington, Wellington

Suratno, T. (2014). The education system in Indonesia at a time of significant changes. *Revue Internationale d'éducation de Sèvres*.

Swain, J. (2018). *A Hybrid Approach to Thematic Analysis in Qualitative Research: Using a Practical Example*. doi:10.4135/9781526435477

Swensson, C. (2017). How similar are the PISA and TIMSS studies? Retrieved from https://ioelondonblog.wordpress.com/2017/12/04/how-similar-are-the-pisa-and-timss-studies/

Syahril, I. (2016). *The Indonesian teacher certification policy: A case study of policy sense-making*. (Doctorate Dissertation). Michigan State University, Michigan.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics (5th ed.)*. Boston: Allyn & Bacon/Pearson Education.

Tagiuri, R. (1968). *The concept of organizational climate*. Boston: Division of Research, Graduate School of Business Administration, Harvard University.

Tan, C. (2014a). Educative Tradition and Islamic Schools in Indonesia *Journal of Arabic and Islamic Studies, 14*, 47-62.

Tan, C. (2014b). Educative tradition and Islamic schools in Indonesia. *Journal of Arabic and Islamic Studies, 14*(2014). 47-62

Tanner, K. (2013). Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education, 12*(3), 322-331.

Tarmidi, & Akbar-Hawadi, R. (2009). Studi kasus: konsep diri akademik siswa tunaprestasi dan terapi perilakuan kognitif pada siswa sekolah dasar. [Case study: academic self-concept of underachiever primary school student and cognitive behavior therapy]. *Anima: Indonesia Psychological Journal, 24*(4), 334-347.

Tarmidi, & Rambe, A. R. R. (2013). Korelasi Antara Dukungan Sosial Orang Tua dan Self-Directed Learning pada Siswa SMA. *Jurnal Psikologi*(Vol 37, No 2 (2010)), 216â□“223. Retrieved from http://jurnal.psikologi.ugm.ac.id/index.php/fpsi/article/view/34

Tarmidi, & Vanita, I. (2008). Hubungan antara kecerdasan emosi dengan kecemasan menghadapi ujian nasional pada siswa SMA swasta. [Relationship between emotional intelligence and exam anxiety on national examination of private senior high school students]. *Jurnal InSight, 6*(1), 66-77.

Tarmidi, & Wulandari, L. H. (2005). Prestasi belajar ditinjau dari persepsi siswa terhadap iklim kelas pada siswa yang mengikuti program percepatan belajar. [Classroom climate and academic achievement of accelerated program of high school students]. *Psikologia, 1*(1).

Tarmidi, & Wulandari, L. H. (2005). Prestasi belajar ditinjau dari persepsi siswa terhadap iklim kelas pada siswa yang mengikuti program percepatan belajar. [Classroom

climate and academic achievement of accelerated program of high school students]. *Psikologia, 1*(1).

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology : combining qualitative and quantitative approaches*. Thousand Oaks: London: Sage.

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks: Sage.

Tayyaba, S. (2012). Rural-urban gaps in academic achievement, schooling conditions, student, and teachers' characteristics in Pakistan. *International Journal of Educational Management, 26*(1), 6-26.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: Falmer.

Teddlie, C., & Reynolds, D. (2001). Countering the Critics: Responses to Recent Criticisms of School Effectiveness Research. *School Effectiveness and School Improvement, 12*(1), 41-82. doi:10.1076/sesi.12.1.41.3458

Teddlie, C., Reynolds, D., & Sammons, P. (2000). The Methodology and Scientific Properties of School Effectiveness Research. In C. Teddlie & D. Reynolds (Eds.), *The International Hanbook of School Effectiveness Research*. London, UK: Falmer Press.

Teddlie, C., & Sammons, P. (2010). Applications of Mixed Methods to The Field of Educational Effectiveness Research. In B. Creemers, L. Kyriakides, & P. Sammons (Eds.), *Methodological Advances in Educational Effectiveness Research* (pp. 115-152). London, UK: Routledge Taylor Francis.

Teddlie, C., & Stringfield, S. (1993). *Schools Make a Difference: Lessons Learned from a 10-Year Study of School Effects*: ERIC.

Teddlie, C., & Tashakkori, A. (2008). *Foundations of mixed methods research : integrating quantitative and qualitative approaches in the social and behavioral sciences*. Los Angeles ; London: Sage.

Ten Have, P. (2007). *Doing conversation analysis: A practical guide* (2nd ed.): Sage.

Teodorović, J. (2011). Classroom and school factors related to student achievement: what works for students? *School Effectiveness and School Improvement, 22*(2), 215-236. doi:10.1080/09243453.2011.575650

Thahir, A., & Hidriyanti, B. (2014). Pengaruh Bimbingan Belajar Terhadap Prestasi Belajar Siswa Pondok Pesantren Madrasah Aliyah Al-Utrujiyyah kota Karang Bandar Lampung. *KONSELI: Jurnal Bimbingan dan Konseling, 1*(2), 63-76.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research, 83*(3), 357-385. doi:10.3102/0034654313483907

Thiele, T., Singleton, A., Pope, D., & Stanistreet, D. (2016). Predicting students' academic performance based on school and socio-demographic characteristics. *Studies in Higher Education*, *41*(8), 1424-1446.

Thomas, F. (2017). 20 Years of TIMSS: Lessons for Indonesia. *Indonesian Research Journal in Education |IRJE, 1*(1). doi:10.22437/irje.v1i1.4333

Thomas, S. (1998). Value-added measures of school effectiveness in the United Kingdom. *Prospects, 28*(1), 91-108. doi:10.1007/bf02737782

Thomas, S. (2001). Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement, 12*(3), 285-322. doi:DOI 10.1076/sesi.12.3.285.3448

Thomas, S., & Mortimore, P. (1996). Comparison of value-added models for secondary-chool effectiveness. *Research Papers in Education, 11*(1), 5-33. doi:10.1080/0267152960110103

Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997). Differential Secondary School Effectiveness: comparing the performance of different pupil groups. *British Educational Research Journal, 23*(4), 451-469. doi:10.1080/0141192970230405

Thomas, S., Smees, R., MacBeath, J., Robertson, P., & Boyd, B. (2000). Valuing Pupils' Views in Scottish Schools. *Educational Research and Evaluation, 6*(4), 281-316. doi:10.1076/edre.6.4.281.6934

Thomas, S., Smees, R., MacBeath, J., Robertson, P., & Boyd, B. (2010). Valuing Pupils' Views in Scottish Schools. *Educational Research and Evaluation, 6*(4), 281-316. doi:10.1076/edre.6.4.281.6934

Thrupp, M. (2001). Recent School Effectiveness Counter-critiques: Problems and possibilities.*British Educational Research Journal. 27*(4), 443-457. doi:10.1080/01411920120071452

Timmermans, A., & Thomas, S. (2014). The impact of student composition on schools' value-added performance: a comparison of seven empirical studies. *School Effectiveness and School Improvement, 26*(3), 487-498. doi:10.1080/09243453.2014.957328

Tobias, J., Wales, J., Syamsulhakim, E., & Suharti. (2014). *Towards Better Education Quality: Indonesia's Promising Path*. London: Overseas Development Institute.

Tobin, M., Lietz, P., Nugroho, D., Vivekanandan, R., & Nyamkhuu, T. (2015). *Using large-scale assessments of students' learning to inform education policy: Insights from the Asia-Pacific region.*. Camberwell: Australian Council for Educational Research

Traag, T., & Van der Velden, R. (2011). Early school-leaving in the Netherlands: the role of family resources, school composition and background characteristics in early school-leaving in lower secondary education. *Irish Educational Studies, 30*(1), 45-62.

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology, 90*(2), 334-349. doi:10.1037/0022-3514.90.2.334

Treagust, D. (2004). The Status of Science Classroom Learning Environments in Indonesian Lower Secondary Schools. *Learning Environments Research, 7*(1), 43-63. doi:10.1023/b:Leri.0000022282.48004.18

Uline, C. (2008). The walls speak: the interplay of quality facilities, school climate, and student achievement. *Journal of Educational Administration, 46*(1), 55-73. doi:10.1108/09578230810849817

UNESCO-UIS. (2012). *A Pace to Learn: Lessons from Research on Learning Environment.* Retrieved from: http://www.uis.unesco.org/Education/Documents/tp9-learning-environments-2012-en2.pdf

UNESCO. (2004). *EFA Global Monitoring Report 2005*. Retrieved from http://unesdoc.unesco.org/images/0013/001373/137333e.pdf

UNESCO. (2014). *Teaching and Leraning: Achieving Quality for all.* Paper presented at the EFA Global Monitoring Report 2013/14, Paris, France.

Urban, V. (1999). Eugene's Story: A Case for Caring. *Educational Leadership, 56*(6), 69-70.

Usher, E., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology, 34*(1), 89-101. doi:10.1016/j.cedpsych.2008.09.002

Valentine, J.., DuBois, D., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*(2), 111-133. doi:DOI 10.1207/s15326985ep3902_3

Van Horn, M. (2003). Assessing the Unit of Measurement for School Climate through Psychometric and Outcome Analyses of the School Climate Survey. *Educational and Psychological Measurement, 63*(6), 1002-1019. doi:10.1177/0013164403251317

Van Houtte, M. (2005). Climate or Culture? A Plea for Conceptual Clarity in School Effectiveness Research. *School Effectiveness and School Improvement, 16*(1), 71-89. doi:10.1080/09243450500113977

Van Houtte, M., & Van Maele, D. (2011). The black box revelation: in search of conceptual clarity regarding climate and culture in school effectiveness research. *Oxford Review of Education, 37*(4), 505-524. doi:10.1080/03054985.2011.595552

Voight, A., Austin, G., & Hanson, T. (2013). *A climate for academic success: How school climate distinguishes schools that are beating the achievement odds.* San Francisco: WestEd.

Wahyudi, & Darrell, F. (2006). School climate in indonesian junior secondary schools. In *Contemporary Approaches to Research on Learning Environments* (pp. 497-516): WORLD SCIENTIFIC.

Wales, J., Magee, A., & Nicolai, S. (2016). *How does political context shape education reforms and their success?*. London: Overseas Development Institute.

Wang, M., Degol, J., & Ye, F. (2015). Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Frontiers in psychology*, *6*, 36.

Wang, M., & Degol, J. (2016). School Climate: a Review of the Construct, Measurement, and Impact on Student Outcomes. *Educational Psychology Review, 28*(2) 1573-336X),

315-352. doi:10.1007/s10648-015-9319-1 %U https://doi.org/10.1007/s10648-015-9319-1

Wang, M., & Eccles, J. (2012). Adolescent Behavioral, Emotional, and Cognitive Engagement Trajectories in School and Their Differential Relations to Educational Success. *Journal of Research on Adolescence, 22*(1), 31-39. doi:10.1111/j.1532-7795.2011.00753.x

Wang, W., Vaillancourt, T., Brittain, H., McDougall, P., Krygsman, A., Smith, D., Cunningham, C., Haltigan, J., Hymel, S. (2014). School Climate, Peer Victimization, and Academic Achievement: Results From a Multi-Informant Study. *School Psychology Quarterly, 29*(3), 360-377. doi:10.1037/spq0000084

Wang, Z., Osterlind, S., & Bergin, D. (2012). Building Mathematics Achievement Models in Four Countries Using Timss 2003. *International Journal of Science and Mathematics Education, 10*(5), 1215-1242. doi:10.1007/s10763-011-9328-6

Warwick, D., & Jatoi, H. (1994). Teacher Gender and Student Achievement in Pakistan. *Comparative Education Review, 38*(3), 377-399.

Watson, D., Emery, C., & Bayliss, P. (2012). *Children's social and emotional wellbeing in schools : a critical perspective*. Bristol: Policy Press.

Weber, G. (1971). Inner-City Children Can Be Taught to Read: Four Successful Schools. CBE Occasional Papers, Number 18.

White, H. (2010). William James's Pragmatism. Ethics and The Individualism of Others. *European Journal of Pragmatism, 2*(II-1).

White, N., La Salle, T., Ashby, J., & Meyers, J. (2014). A Brief Measure of Adolescent Perceptions of School Climate. *School Psychology Quarterly, 29*(3), 349-359. doi:10.1037/spq0000075

Wiberg, M. (2019). The relationship between TIMSS mathematics achievements, grades, and national test scores. *Education Inquiry*, 1-16. doi:10.1080/20004508.2019.1579626

Widayanti, C., & Siswati, S. (2009). Fenomena bullying di sekolah dasar negeri di semarang: sebuah studi deskriptif. *Jurnal Psikologi Undip*. 5(2). no page number.

Wihardit, K. (2010). Pendidikan multikultural: suatu konsep, pendekatan dan solusi. *Jurnal Pendidikan, 11*(2), 96-105.

Wijaya, A. (2017). The difficulties of Indonesian fourth graders in learning fractions: An early exploration of TIMSS 2015 results. *1868*(1), 050027. doi:10.1063/1.4995154

Willms, J., & Raudenbush, S. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26*(3), 209-232.

Wilson, M., & Greenhill, A. (2004). Theory and action for emancipation: elements of a critical realist approach. In *Information Systems Research* (pp. 667-674). Boston: Springer.

Winter, J. (1987). *A qualitative study of teachers' perceptions of school climate utilizing the interview method.* (PhD). Iowa State University, Iowa.

Withall, J. (1949). The Development of a Technique for the Measurement of Social-Emotional Climate in Classrooms. *The Journal of Experimental Education, 17*(3), 347-361. doi:10.1080/00220973.1949.11010391

Withall, J. (1969). Evaluation of Classroom Climate. *Childhood Education, 45*(7), 403-408. doi:10.1080/00094056.1969.10729475

Wiwoho, L. (2018). UNBK dan Cara Berpikir Kritis. *Kompas*. Retrieved from https://edukasi.kompas.com/read/2018/04/30/07520021/unbk-dan-cara-berpikir-kritis?page=all

Wolcott, H. (1994). *Transforming qualitative data : description, analysis, and interpretation*. Thousand Oaks: Sage.

World Bank. (2014a). Data - Indonesia. Retrieved from http://data.worldbank.org/country/indonesia

World Bank. (2014b). *Tertiary Education in Indonesia: Directions for Policy*. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2014/07/15/000333037_20140715141201/Rendered/PDF/893010WP0P12940olicy0Paper0June2014.pdf

World Bank. (2014c). World Bank and Education in Indonesia. Retrieved from http://www.worldbank.org/en/country/indonesia/brief/world-bank-and-education-in-indonesia

World Bank. (2019). School Enrollment, secondary (%net) - Indonesia. Retrieved from https://data.worldbank.org/indicator/SE.SEC.NENR?locations=ID.

Worrell, F. (2007). Ethnic identity, academic achievement, and global self-concept in four groups of academically talented adolescents. *Gifted Child Quarterly*, *51*(1), 23-38.

Yang, C., Bear, G., Chen, F., Zhang, W., Blank, J., & Huang, X. (2013). Students' perceptions of school climate in the U.S. and China. *School Psychology Quarterly, 28*(1), 7-24. doi:10.1037/spq0000002

Yeom, M., Acedo, C., Utomo, E., & Yeom, M. (2002). The reform of secondary education in indonesia during the 1990s: Basic education expansion and quality improvement through curriculum decentralization. *Asia Pacific Education Review*, *3*(1), 56-68. doi:10.1007/bf03024921

Yeung, A., & Lee, F. (1999). Self-concept of high school students in China: Confirmatory factor analysis of longitudinal data. *Educational and Psychological Measurement, 59*(3), 431-450. doi:10.1177/00131649921969965

Yin, R. (2014). *Case study research : design and methods* (5th ed.). Los Angeles: SAGE.

Young, D. (1998). Rural and urban differences in student achievement in science and mathematics: A multilevel analysis. *School Effectiveness and School Improvement, 9*(4), 386-418. doi:Doi 10.1080/0924345980090403

Yu, G., & Thomas, S. (2008). Exploring school effects across southern and eastern African school systems and in Tanzania. *Assessment in Education: Principles, Policy & Practice, 15*(3), 283-305. doi:10.1080/09695940802417525

Zerin, S. (2009). Classroom management: Seating arrangements in ESL classroom.

Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *Annals of translational medicine, 4*(1), 9. doi:10.3978/j.issn.2305-5839.2015.12.38

Zimmerman, B. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemp Educ Psychol, 25*(1), 82-91. doi:10.1006/ceps.1999.1016

Zins, J., & Elias, M. (2007). Social and emotional learning: Promoting the development of all students. *Journal of Educational and Psychological consultation, 17* (2-3), 233-255.

Zuhdi, M. (2006). Modernization of Indonesian Islamic schools' curricula, 1945–2003. *International Journal of Inclusive Education, 10*(4-5), 415-427. doi:10.1080/13603110500430765

Zullig, K., Koopman, T., Patton, J., & Ubbes, V. (2010). School climate: Historical review, instrument development, and school assessment. *Journal of psychoeducational assessment, 28*(2), 139-152.

Zullig, K., Huebner, E., & Patton, J. (2011). Relationships among School Climate Domains and School Satisfaction. *Psychology in the Schools, 48*(2), 133-145. doi:10.1002/pits.20532

Appendices

Appendix 1: Ethics approval

Research and Enterprise Development University of Bristol

Application Review Logout

ID	Name	Faculty	Department	Supervisor
19581	Mr Tarmidi Dadeh	Faculty of Social Sciences and Law	Graduate School of Education	Professor Sally Thomas

Status
Signed off
Date added
March 17, 2015
Signed off date
Sept. 3, 2015
Is this a student project?
Postgraduate Phd
Project title
School climate and student learning outcomes; A mixed-methods study of junior secondary school in Indonesia
Estimated start date
April 10, 2015
Duration (months)
20
Project outline
This study aims to explore the effect of school climate on students' learning outcomes both academic and self-beliefs in the context of Indonesian. The study is informed by several issues in the academic achievement of Indonesian students that always reported as lower ranked among other countries who participated in international educational assessments like PISA and TIMSS. The study draws on school effectiveness research (SER) theories, conceptual models, methodologies, and approach. Particularly uses the dynamic model of educational effectiveness given that, evidence from the research field links the important of school climate as one factor that can give an explanation of quality of education. The study is divided into two phases. First, quantitative multilevel analysis of the TIMSS 2011 dataset will be conducted to develop general understanding of the relationship between school climate and learning outcomes. The second phase, four schools will be selected to be involved in the case study analysis to explore school members' views of their climate and its effect on student learning in more depth. The school members that will participate in this study including head teachers, teachers, and students. The headmaster will be interviewed by the researcher asking about school policy in creating school climate to enhance student learning. Next, a group of teacher in each school will be involved in the Focus Group Discussion (FGD) session to discuss how school policy in creating school climate affect their teaching-learning process in the classroom. Finally, a group of student also invited to involve in FGD to explore how they feel about their school climate and how its effect on their learning.

The interview and FGD will be electronically recorded, transcribed, and analysed. Also, these qualitative data will be integrated with quantitative data to build an explanation of the study.

Supporting information

I will conduct my research in my own country. Also, I will access the names of schools on TIMSS 2011 datasets from a confidential government source and have permission for this.

Files

Letter of Access ethics/19581/email-from-moec.pdf (109.0 KB added on Aug. 26, 2015)
Peer Review ethics/19581/ethicsformrev.pdf (1.7 MB added on Aug. 26, 2015)
Consent Form ethics/19581/consent_form_teachers-and-head-teacher_tOQQPiu.docx (76.2 KB added on June 9, 2015)
Participant Information Sheet ethics/19581/participant-information-sheet_w4Mk4b5.docx (56.1 KB added on June 9, 2015)
Consent Form ethics/19581/consent_form_student_DXf6QQe.docx (82.4 KB added on June 9, 2015)

L1. Does your research involve any of the following?

- Medical Devices, ionising radiation, drugs, placebos or other substances to be administered to participants.
- Human Blood or Tissue Samples (Tissue means any relevant material consisting of or including cells - for definition of 'relevant material', please see the Human Tissue Authority website at http://www.hta.gov.uk/ - this link opens in a new window).
- Adults (over 16) who lack capacity to consent for themselves including participants, who will be retained in the study following loss of Capacity.
- Recruiting or using client data from NHS patients, nursing home/independent hospital/clinic or medical agency patients, users of social care services or prisoners. For more details on definitions please see 'Does my project require review': http://www.nres.nhs.uk/applications/approval-requirements/ethical-review-requirements/requirements-for-ethical-review-under-legislation/ (this link opens in a new window).

No

L2. Does your research involve any of the following?

- Animals (either use or observation)
- Has or will your research be submitted to another ethics committee? (If so please provide details of the committee and dates (submission/approval/provisional approval etc.)

No

L3. Does your research involve any of the following?

- Working or travelling overseas
- Trials outside the UK
- Pregnant research subjects
- Conception/Contraception
- Children under 5
- More than 1500 research subjects
- Genetic Engineering

Yes

- Hepatitis/CJD/HIV & AIDS related research

1. Does the research involve human participants?

If you answered No, please go to question 2.

Yes

1a. Does the research involve participants who are particularly vulnerable or unable to give informed consent?

Examples of vulnerable participants or those unable to give informed consent are children, people with learning difficulties, patients, people experiencing emotional distress or mental illness, people living in care or nursing homes, and people recruited through self-help groups, participants in a dependent or unequal relationship with the researcher(s) or research supervisor.

Yes

1b. Will it be necessary for participants to take part without their knowledge and consent at the time?

Examples include the covert observation of people

No

1c. Will the research involve actively deceiving participants?

Examples include deliberately falsely informing participants, withholding information from participants or misleading participants in such a way that they are likely to object or show unease when debriefed about the study.

No

1d. Will the research involve discussion or collection of information on sensitive topics?

Sensitive topics under the Data Protection Act 1998 include:

- The racial or ethnic origin of the data subject;
- Their religious beliefs or other beliefs of a similar nature;
- Whether they are a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992);
- Their physical or mental health or condition;
- Their sexual life;
- Their commission or alleged commission by them of any offence;
- Any proceedings for any offence committed or alleged to have been committed by them, the disposal of such proceedings or the sentence of any court in such proceedings.

Yes

If the research is in relation to any of the sensitive topics listed under the DPA 1998 then the legal issue requiring such scrutiny in such cases that 'explicit consent' must be obtained.

1e. Does the research involve invasive procedures?

Invasive procedures may include:

- Administration of drugs placebos, or other substances (e.g., drinks, foods, food or drink constituents, dietary supplements) to study participants;
- Biological samples from participants be obtained;
- Pain or more than mild discomfort likely to result from the study.

No

1f. Does the research involve scans or x-rays of research

participants?	
1g. Does the research involve photographs, videoing, recording or similar of research participants?	No
1h. Will financial inducement (other than reasonable expenses and compensation for time) be offered?	No
1i. Will the study involve the use or storage of information about living people whose personal identity could be discovered from that information?	No
1j. Does the study risk causing psychological stress or anxiety or other harm or negative consequences beyond that normally encountered by the participants in their life outside research?	No
2. Will the research involve politically and culturally sensitive funding sources? Examples include the defence sector, projects with potential environmental effects and other internationally regulated or protected industries. For more information, please follow the link to the 'Research Governance and Integrity Policy': **http://www.bris.ac.uk/red/support/governance/RGI.pdf** (this link opens in a new window).	No
3. Will the research involve politically, culturally or socially sensitive topics? For more information, please follow the link to the **Faculty of Arts Ethics Committee Guidance Note (PDF 78kb)** (this link opens in a new window).	No
Supporting information (up to approximately 300 words) Please provide any additional information in relation to your study such as adhering to a particular SOP or confirming if your study is a service evaluation/audit as opposed to research.	I will conduct my research in my own country. Also, I will access the names of schools on TIMSS 2011 datasets from a confidential government source and have permission for this.

Flesch Reading Ease (http://www.readabilityformulas.com/free-readability-formula-tests.php)

Material	**Reading ease score**

Faculty guidance and ethics application

http://www.bris.ac.uk/fssl/current-staff/researchethics/

Users guides Submitter Approver Reviewer Signatory

Updated 20 December 2012 by Research and Enterprise Development | Feedback
University of Bristol, Senate House, Tyndall Avenue, Bristol BS8 1TH, UK. Tel: +44 (0)117 928 8676

Appendix 2: Letter of Acceses

University of BRISTOL

Tarmidi Dadeh <td13465@bristol.ac.uk>

Izin Penggunaan Data Indonesia pada TIMSS 2011

5 messages

Tarmidi Dadeh <t.dadeh@bristol.ac.uk> 27 August 2014 at 09:18
Reply-To: t.dadeh@bristol.ac.uk
To: [redacted]

Yth Bapak Prof. Nizam
Kepala Pusat Penilaian Pendidikan Balitbang Kemdikbud RI

Dengan hormat,
Mohon maaf sebelumnya Prof, saya mendapatkan email ini dari web OECD. Selain itu juga atas saran dari supervisor saya sewaktu mengambil Master Psikologi Pendidikan di UI (Ibu Reni Hawadi), bahwa Bapak orang yang tepat untuk dihubungi tentang studi TIMSS di Indonesia. Saya juga membaca informasi dari studi TIMSS dan PISA bahwa bapak adalah National Research Manager untuk kegiatan-kegiatan tersebut.
Perkenalkan saya dosen di Fakultas Psikologi Univ. Sumatera Utara Medan. Saat ini sedang menjalani Program PhD di Graduate School of Education, University of Bristol, UK. Berkaitan dengan penelitian yang akan saya lakukan yang salah satunya menggunakan data pendidikan Indonesia pada studi TIMSS, ada beberapa hal yang ingin saya mohon bantuan bapak.
1. Saya mohon izin menggunakan data pendidikan Indonesia pada TIMSS untuk saya gunakan dalam riset PhD saya. Jika sudah mendapat izin prinsip, selanjutnya akan saya susul permohonan izin resmi dari University of Bristol.
2. Saya membutuhkan data yang lebih detail mengenai data-data Indonesia yang tidak tercantum di TIMSS (seperti data2 sekolah yang berperan serta, data sekolah negeri/swasta, umum/madrasah, kepemilikan yang khas Indonesia untuk menggambarkan keadaan sosial ekonomi siswa).
Adapun riset yang akan saya lakukan bertemakan School Effectiveness dan aspek sosial-psikologis yang mempengaruhi (iklim sekolah, konsep diri siswa, motivasi, dan aspek non akademis lainnya).
Secara pribadi saya berharap bahwa penelitian yang akan saya lakukan bisa bermanfaat untuk perkembangan pendidikan di Indonesia.
Besar sekali harapan saya untuk mendapatkan informasi lebih lanjut mengani hal ini. Demikian saya sampaikan, semoga Bapak berkenan. Terimakasih saya ucapkan sebelumnya.

Salam hormat,
Tarmidi

Tarmidi H Dadeh
PhD Student
Graduate School of Education
University of Bristol, United Kingdom

--
Tarmidi H Dadeh

PhD Student
Graduate School of Education
University of Bristol, United Kingdom

Tarmidi Dadeh <t.dadeh@bristol.ac.uk> 15 September 2014 at 12:28
Reply-To: t.dadeh@bristol.ac.uk
To: [redacted]

Yth Bapak Prof. Nizam
Kepala Pusat Penilaian Pendidikan Balitbang Kemdikbud RI

Dengan hormat,
Mohon maaf sebelumnya Prof, saya mendapatkan email ini dari web OECD. Selain itu juga atas saran dari supervisor saya sewaktu mengambil Master Psikologi Pendidikan di UI (Ibu Reni Hawadi), bahwa Bapak orang yang tepat untuk dihubungi tentang studi TIMSS di Indonesia. Saya juga membaca informasi dari studi TIMSS dan PISA bahwa bapak adalah National Research Manager untuk kegiatan-kegiatan tersebut.
Perkenalkan saya dosen di Fakultas Psikologi Univ. Sumatera Utara Medan. Saat ini sedang menjalani Program PhD di Graduate School of Education, University of Bristol, UK. Berkaitan dengan penelitian yang akan saya lakukan yang salah satunya menggunakan data pendidikan Indonesia pada studi TIMSS, ada beberapa hal yang ingin saya mohon bantuan bapak.
1. Saya mohon izin menggunakan data pendidikan Indonesia pada TIMSS untuk saya gunakan dalam riset PhD saya. Jika sudah mendapat izin prinsip, selanjutnya akan saya susul permohonan izin resmi dari University of Bristol.
2. Saya membutuhkan data yang lebih detail mengenai data-data Indonesia yang tidak tercantum di TIMSS (seperti data2 sekolah yang berperan serta, data sekolah negeri/swasta, umum/madrasah, kepemilikan yang khas Indonesia untuk menggambarkan keadaan sosial ekonomi siswa).
Adapun riset yang akan saya lakukan bertemakan School Effectiveness dan aspek sosial-psikologis yang mempengaruhi (iklim sekolah, konsep diri siswa, motivasi, dan aspek non akademis lainnya).
Secara pribadi saya berharap bahwa penelitian yang akan saya lakukan bisa bermanfaat untuk perkembangan pendidikan di Indonesia.
Besar sekali harapan saya untuk mendapatkan informasi lebih lanjut mengani hal ini. Demikian saya sampaikan, semoga Bapak berkenan. Terimakasih saya ucapkan sebelumnya.

Salam hormat,
Tarmidi

--
Tarmidi H Dadeh

PhD Student
Graduate School of Education
University of Bristol, United Kingdom

Nizam [redacted] 17 September 2014 at 10:17
To: t dadeh <t.dadeh@bristol.ac.uk>
Cc: [redacted]

Saudara Tarmidi Ysh.,
Pada prinsipnya data hasil survey yang sifatnya bukan rahasia dapat dipergunakan untuk penelitian. Kami mendukung penelitian Saudara di U Bristol. Kami juga sedang melakukan kajian-kajian lanjutan dari hasil PISA dan TIMSS. Teman-teman peneliti di Puspendik akan senang hati bekerja sama dalam penelitian ini dengan Saudara. Silahkan berkoordinasi dengan Kabid Pendataan dan Informasi, Dr. Mahdiansyah Makmur,
Wassalam
Nizam

From: "Tarmidi Dadeh" <t.dadeh@bristol.ac.uk>
To: [redacted]
Sent: Monday, September 15, 2014 6:28:59 PM
Subject: Fwd: Izin Penggunaan Data Indonesia pada TIMSS 2011
[Quoted text hidden]

Tarmidi Dadeh <t.dadeh@bristol.ac.uk> 17 September 2014 at 14:54
Reply-To: t.dadeh@bristol.ac.uk
To: Nizam [redacted]

Prof Nizam ysh, terimakasih atas izin dan dukungannya

Salam

Tarmidi H Dadeh

PhD Student
Graduate School of Education
University of Bristol, United Kingdom

[Quoted text hidden]

Mahdiansyah Makmur [redacted] 22 September 2014 at 06:36
To: t.dadeh@bristol.ac.uk

Aswrwb p. Tarmidi,
berikut kami sampaikan data sekolah sampel TIMMS 2011, sebagaimana yang diminta.

Wass,
Mahdiansyah

---

daftar sekolah main survey timss 2011.xlsx
102K

Appendix 3: Peer review ethics form

GSoE RESEARCH ETHICS FORM

Name: Tarmidi H. Dadeh
Proposed research project: School climate and students' learning outcomes: A mixed method study of junior secondary school in Indonesia

Proposed funder: Directorate General of Higher Education, Ministry of Education and Culture, Republic of Indonesia

Discussant for the ethics meeting: Lei Zhang (Fellow student)

Name of supervisors: Professor Sally Thomas & Dr Guoxing Yu

Has your supervisor seen this submitted draft of your ethics application? Y

Project Outline:

This study aims to explore the effect of school climate on students' learning outcomes both academic and self-beliefs in the context of Indonesian. The study is informed by several issues in the academic achievement of Indonesian students that always reported as lower ranked among other countries who participated in international educational assessments like PISA and TIMSS. The study draws on school effectiveness research (SER) theories, conceptual models, methodologies, and approach. Particularly uses the dynamic model of educational effectiveness given that, evidence from the research field links the important of school climate as one factor that can give an explanation of quality of education.

The study is divided into two phases. First, quantitative multilevel analysis of the TIMSS 2011 dataset will be conducted to develop general understanding of the relationship between school climate and learning outcomes. The second phase, four schools will be selected to be involved in the case study analysis to explore school members' views of their climate and its effect on student learning in more depth. The school members that will participate in this study including head teachers, teachers, and students. The headmaster will be interviewed by the researcher asking about school policy in creating school climate to enhance student learning.

Next, a group of teacher in each school will be involved in the Focus Group Discussion (FGD) session to discuss how school policy in creating school climate affect their teaching-learning process in the classroom. Finally, a group of student also invited to involve in FGD to explore how they feel about their school climate and how its effect on their learning. The interview and FGD will be electronically recorded, transcribed, and analysed. Also, these qualitative data will be integrated with quantitative data to build an explanation of the study.

Ethical issues discussed and decisions taken

The meeting discussed the issues that might come up in this research design as suggested by Graduate School of Education Ethic guidelines. Since this research uses secondary data from TIMSS (public data), the ethical issues in the discussion are mostly cover for the qualitative phase. The following sections are ethical issues that should be considered in this research.

Data Protection

This research use two types of data, first is quantitative data taken from TIMSS 2011 dataset, and the second is qualitative data that will be obtained from interviews and FGDs. For that reason, the researcher must comply with those types of data. For the first data, because the data is available to the public, there is not necessary to protect the data securely. However, the researcher must consider how the result will be storage securely because it is still sensitive to school performance. Therefore, the result will be saved in researcher's password protected computer files. Since the TIMSS data will be used to select school to involve in the next research phase, written permission from Ministry of Education and Culture (MOEC) of the Republic of Indonesia or local government authorities is needed.

Additionally, in the informed consent form, the researcher and researcher's supervisor email address will be included to enquire further or to make any complains that may arise in

the research process. Moreover, in line with BERA (2011) guiding principle concerning to unexpected effect which can rise from the research, the schools will notify of the possibility that the findings will not certainly present the school in a positive side.

Researcher's position

The researcher should pay attention to minimizing power relation when working with the participants. Therefore, the researcher needs to maintain the relationship with all participants as a partner in working together to make school as a better place to learn. However, there are still conditions when power relationship is apparent. For example, students or teachers may participate in the research process because they are pushed by the school. For that reason, the researcher needs to inform clearly that all participants have the right to withdraw at any time in the research activities. Also, student, for instance, may not want to follow research activities, but their school or teacher might push them to do so. Therefore, the researcher needs to consider to make sure that students do not experience any distress or discomfort.

Additionally, researcher bias is also an important aspect that will be always considered, and the researcher approach in doing research will actively seek to avoid bias and prejudices. The bias may perhaps have an impact when the researcher analyses the data. For example, researcher's identity as a Moslem, when analysing data from the Islamic school may possibly cause bias in some cases. However, it is impossible to avoid the pitfalls of the researcher identity. Therefore, the researcher should keep the position as an outsider. Although this research involves selecting Islamic schools as case study schools, detail or specific given religious views will not sought or be of interest in this research. However, there may be some links between religious views and school climate expressed by some stakeholders. It is not anticipated that this viewpoints if expressed in any way will be sensitive or harmful to those involved in the study.

Other moral concern

School identification as an effective or less effective in term of their climate may risk negative effect on particular schools. Therefore, the researcher should not inform any school regarding a label such as good or bad (that may be used in the methodology of selecting the sample) but rather treat all school context as an opportunity to improve schools as a better place for delivering teaching-learning process.

Also, access to the participants might affect school activities though this will be discussed with head-teachers. Thus, the researcher need to have a positive and respectful attitude to every school and have to comply with the school rules (including the time when interview and FGD can be carried out) during the research process. Finally, since the quantitative analysis of TIMSS data from the first phase will inform the second phase of the study (case study), it is essential to get written permission from MOEC of Indonesia or local government authorities to identify the result of individual schools. This permission was granted to access the school data from Pusat Penilaian Pendidikan (National Assessment Centre of Education - NACE) MOEC. Please refer to email from nizam@kemdikbud.go.id – Chief of NACE on 17th September 2014, and mahdiansyah2007@gmail.com – Head of Data and Information at NACE on 22 September 2014.

Signed by researcher: [redacted] (Tarmidi Dadeh)

Signed by discussant: [redacted] (Lei Zhang)

Date: 12 March 2015

Revised on 24 August 2015

Appendix 4: Consent form

University of BRISTOL

CONSENT FORM

Judul Penelitian *Research Title*:
Iklim Sekolah dan Hasil belajar: Penelitian pada siswa sekolah menengah pertama di Indonesia dengan menggunakan mixed-method
School Climate and Students' Learning Outcomes: A Mixed method study of the Junior Secondary School in Indonesia.

Peneliti *Researcher*:
Tarmidi Dadeh
Graduate School of Education, University of Bristol, UK

Mohon beri tanda
Please initial box

1. Saya menyatakan bahwa saya telah membaca dan memahami lembar informasi mengenai penelitian ini dan mendapat kesempatan untuk bertanya.
   I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.

2. Saya memahami bahwa keikutsertaan dalam penelitian ini bersifat sukarela dan saya bisa menarik diri kapan saja tanpa memberikan alasan apapun.
   I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.

3. Saya bersedia untuk berpartisipasi dalam penelitian ini.
   I agree to take part in the above study

Mohon beri tanda
Please initial box

Ya	Tidak
Yes	No

4. Saya bersedia bahwa interview / focus group ini direkam dengan menggunakan perekam suara.
   I agree to the interview / focus group being audio recorded

5. Saya setuju bahwa penggunaan pernyataan saya bersifat anonim (tidak mencantumkan nama) pada saat mempublikasikan hasil-hasil penelitian ini.
   I agree to the use of selected anonymised quotes in publications.

6. Saya setuju bahwa data yang diperoleh dalam penelitian ini disimpan pada dokumen yang dijaga kerahasiaannya dan dapat digunakan untuk penelitian selanjutnya.
   I agree that my data gathered in this study may be stored (after it has been anonymised) in a protected electronic document and may be used for future research.

Nama partisipan *Name of Participant* **Tanggal** *Date* **Tanda Tangan** *Signature*

Nama Peneliti *Name of Researcher* **Tanggal** *Date* **Tanda Tangan** *Signature*

University of BRISTOL

CONSENT FORM
Izin orangtua *Parental Permission*

Judul Penelitian *Research Title*:
Iklim Sekolah dan Hasil belajar: Penelitian pada siswa sekolah menengah pertama di Indonesia dengan menggunakan mixed-method
School Climate and Students' Learning Outcomes: A Mixed-method study of the Junior Secondary School in Indonesia.

Peneliti *Researcher*:
Tarmidi Dadeh
Graduate School of Education, University of Bristol, UK

Mohon beri tanda
Please initial box

1. **Saya menyatakan bahwa saya telah membaca dan memahami lembar informasi mengenai penelitian ini dan mendapat kesempatan untuk bertanya.**
   I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.

2. **Saya memahami bahwa keikutsertaan anak saya dalam penelitian ini bersifat sukarela dan anak saya bisa menarik diri kapan saja tanpa memberikan alasan apapun.**
   I understand that my child's participation is voluntary and that my child is free to withdraw at any time, without giving reason.

3. **Saya menyatakan bahwa saya bersedia bahwa anak saya berpartisipasi dalam penelitian ini.**
   I give permission for my child to participate in the above study.

Mohon beri tanda
Please initial box

Ya Yes	**Tidak** No

4. **Saya bersedia bahwa interview / focus group anak saya direkam dengan menggunakan perekam suara.**
   I agree to the interview / focus group of my child being audio recorded

5. **Saya setuju bahwa penggunaan pernyataan anak saya bersifat anonim (tidak mencantumkan nama) pada saat mempublikasikan hasil-hasil penelitian ini.**
   I agree to the use of selected anonymised quotes of my child's comments in publications.

6. **Saya setuju bahwa data anak saya yang diperoleh dalam penelitian ini disimpan pada dokumen yang dijaga kerahasiaannya dan dapat digunakan untuk penelitian selanjutnya.**
   I agree that my child data gathered in this study may be stored (after it has been anonymised) in a protected document and may be used for future research.

Nama anak *Child's name* **Tanggal** *Date* **Tanda Tangan** *Signature*

Nama orangtua/wali *Parent's name* **Tanggal** *Date* **Tanda Tangan** *Signature*

Hubungan dengan anak (contoh: ayah, ibu, wali) *Relationship to child (e.g. mother, father, guardian):* _____

Nama Peneliti *Name of Researcher* **Tanggal** *Date* **Tanda Tangan** *Signature*

Appendix 5: Participant Information sheet

University of BRISTOL

PARTICIPANT INFORMATION SHEET

Hello my name is Tarmidi Dadeh, I am a PhD student at the Graduate School of Education, University of Bristol, United Kingdom. I would like to invite you to contribute in this research project which forms part of my PhD research. Before you decide whether you want to take part, it is essential for you to understand why the research is being conducted and what your participation will involve. Please read the following information carefully and discuss it with others if you wish. Ask me if there is anything that is not clear or if you would like more information.

What is the Research objectives?
The aim of the study is to explore the nature of the school and classroom climate in Indonesia and the effect of those climates on students' learning outcomes both academic and self-beliefs in the context of Indonesian junior secondary schools. For that reason, obtaining perspectives of teachers, students, as well as head teachers on their experience within the school and classroom is essential.

Why you should participate?
Your participation will provide important data for this research and enable you to share your views and experiences regarding school climate and classroom climate in your school and more generally.

Do I have to participate?
Your participation is voluntary. You do not have to take part. You should not agree to take part in this research until you have had all your questions answered satisfactorily.

What will happen to me if I participate?
If you agree to take part, you will be given this information sheet to keep and will be requested to sign a consent form. Then, I will call you to discuss the interview or focus group discussion (FGD) procedure. The interview and the FGD will take roughly one hour and be based on the topic guide, but it is planned to be flexible. These processes will be recorded, subject to your agreement. Recordings of these interview and FGD will be deleted after transcription. Even if you have decided to participate, you are still free to stop your participation at any time and to have research data/information relating to you withdrawn without giving any reason.

What are the possible risks of participation?
There are no foreseeable risks in participating in the study. The main disadvantage to taking part in the study is that you will be donating around an hour of your time to take part. It is possible that you may find answering some of the questions challenging. This is unlikely but if it remained to occur the interview and FGD could be terminated at any time.

What are the possible benefits of participation?
Participating in this interview or FGD may not benefit you personally. However, the information from this study will help to inform policy makers about how school climate affect student learning outcomes. Additionally, I will offer the school with a summary of a final report of this research.

Will my participation be saved confidential?
This research is conducted by following the ethical guidelines of carrying out research in social science. All data for analysis will be anonymised. In reporting on the research findings, I will not reveal the names of any participants or the school. At all times there will be no possibility of you as

individuals being linked with the data and the data is kept on password-protected computer files. Also, all recordings of data on audio recorder will be deleted after transcription.

What will happen to the results of the study?
I will produce a final report summarising the main findings, which will be sent to school. I also plan to disseminate the research findings through publication and conferences.

What if I have further questions or if something goes wrong?
If you have any questions or need more information about this study, please contact me using the following contact details:

Tarmidi Dadeh
Graduate School of Education
University of Bristol
Bristol BS8 1JA, United Kingdom
Email: t.dadeh@bristol.ac.uk

If you wish to make a complaint about the conduct of the study you can contact my main supervisor at the University of Bristol using the details below for further advice and information:

Professor Sally Thomas
Graduate School of Education
University of Bristol
35 Berkeley Square
Bristol BS8 1JA
Email: s.thomas@bristol.ac.uk

~~~~~~~~~~~~~~~~

"Thank You for reading this information sheet and for considering participation in this research"

~~~~~~~~~~~~~~~~

Appendix 6: TIMSS variables used in the study

School climate

STUDENT LEVEL

Factors		Item's number	Item description
Thapa et al. (2013)	**TIMSS**		
Safety	BSBGSBS Students Bullied at School	BSBG13A	1) I was made fun of or called names
		BSBG13B	2) I was left out of games or activities by other students
		BSBG13C	3) Someone spread lies about me
		BSBG13D	4) Something was stolen from me
		BSBG13E	5) I was hit or hurt by other student(s) (e.g., shoving, hitting, kicking)
		BSBG13F	6) I was made to do things I didn't want to do by other students
Teaching and Learning	BSBGEML Students Engaged in Mathematics Lessons	BSBG15A	1) I know what my teacher expects me to do
		BSBG15B	2) I think of things not related to the lesson
		BSBG15C	3) My teacher is easy to understand
		BSBG15D	4) I am interested in what my teacher says
		BSBG15E	5) My teacher gives me interesting things to do
Institutional environment	NA	BSBG12A	1) I like being in school
		BSBG12C	3) I feel like I belong at this school

TEACHER LEVEL

Safety	BTBGSOS Safe and Orderly School	BTBG07A	1) This school is located in a safe neighbourhood
		BTBG07B	2) I feel safe at this school
		BTBG07C	3) This school's security policies and practices are sufficient
		BTBG07D	4) The students behave in an orderly manner
		BTBG07E	5) The students are respectful of the teachers
Teaching learning	BTBGEAS Emphasis on Academic Success (teacher report)	BTBG06B	1) Teachers' understanding of the school's curricular goals
		BTBG06C	2) Teachers' degree of success in implementing the school's curriculum
		BTBG06D	3) Teachers' expectations for student achievement
		BTBG06E	4) Parental support for student achievement
		BTBG06H	5) Students' desire to do well in school
	BTBGIES Instruction to Engage Students in Learning	BTBG14A	1) Summarise what students should have learned from the lesson*
		BTBG14C	2) Use questioning to elicit reasons and explanations
		BTBG14D	3) Encourage all students to improve their performance
		BTBG14E	4) Praise students for good effort
	BTBGCTM	BTBM18A	1) Answer students' questions about mathematics
		BTBM18B	2) Show students a variety of problem-solving strategies
		BTBM18C	3) Provide challenging tasks for capable students

	Confidence in Teaching Mathematics	BTBM18D	4) Adapt my teaching to engage students' interest
		BTBM18E	5) Help students appreciate the value of learning mathematics
Relationship	BTBGCIT Collaborate to Improve Teaching	BTBG10A	1) Discuss how to teach a particular topic
		BTBG10B	2) Collaborate in planning and preparing instructional materials
		BTBG10C	3) Share what I have learned about my teaching experiences
		BTBG10D	4) Visit another classroom to learn more about teaching
		BTBG10E	5) Work together to try out new ideas
Institutional environment	BTBGTWC Teacher Working Conditions	BTBG08A	1) The school building needs significant repair
		BTBG08B	2) Classrooms are overcrowded
		BTBG08C	3) Teachers have too many teaching hours
		BTBG08D	4) Teachers do not have adequate workspace (e.g., for preparation, collaboration, or meeting with students
		BTBG08E	5) Teachers do not have adequate instructional materials and supplies
	BTBGTCS Teacher Career Satisfaction	BTBG11A	1) I am content with my profession as a teacher
		BTBG11B	2) I am satisfied with being a teacher at this school
		BTBG11C*	3) I had more enthusiasm when I began teaching than I have now*
		BTBG11D	4) I do important work as a teacher
		BTBG11E	5) I plan to continue as a teacher for as long as I can
		BTBG11F*	6) I am frustrated as a teacher*

SCHOOL LEVEL

Teaching learning	BCBGEAS School Emphasis on Academic Success–Principal Reports	BCBG11B	1) Teachers' understanding of the school's curricular goals
		BCBG11C	2) Teachers' degree of success in implementing the school's curriculum
		BCBG11D	3) Teachers' expectations for student achievement
		BCBG11E	4) Parental support for student achievement
		BCBG11H	5) Students' desire to do well in school
Institutional environment	BCBGMRS Instruction Affected by Mathematics Resource Shortages (principal report)		**A. General School Resources**
		BCBG09AA	1) Instructional materials (e.g., textbooks)
		BCBG09AB	2) Supplies (e.g., papers, pencils
		BCBG09AC	3) School buildings and grounds
		BCBG09AD	4) Heating/cooling and lighting systems
		BCBG09AE	5) Instructional space (e.g., classrooms)
		BCBG09AF	6) Technologically competent staff
			B. Resources for Mathematics Instruction
		BCBG09BA	1) Teachers with a specialisation in mathematics
		BCBG09BB	2) Computers for mathematics instruction
		BCBG09BC	3) Computer software for mathematics instruction
		BCBG09BD	4) Library materials relevant to mathematics instruction
		BCBG09BE	5) Audio-visual resources for mathematics instruction
		BCBG09BF	6) Calculators for mathematics instruction
Safety	BCBGDAS	BCBG12AA	1) Arriving late at school
		BCBG12AB	2) Absenteeism (i.e., unjustified absences)

	School Discipline and Safety	BCBG12AC	3) Classroom disturbance
		BCBG12AD	4) Cheating
		BCBG12AE	5) Profanity
		BCBG12AF	6) Vandalism
		BCBG12AG	7) Theft
		BCBG12AH	8) Intimidation or verbal abuse among students (including texting, emailing, etc.)
		BCBG12AI	9) Physical injury to other students
		BCBG12AJ	10) Intimidation or verbal abuse of teachers or staff (including texting, emailing, etc.)
		BCBG12AK	11) Physical injury to teachers or staff
Leadership	NA	BCBG17A	1) Promoting the school's educational vision or goals
		BCBG17B	2) Developing the school's curricular and educational goals
		BCBG17C	3) Monitoring teachers' implementation of the school's educational goals in their teaching
		BCBG17D	4) Monitoring students' learning progress to ensure that the school's educational goals are reached
		BCBG17E	5) Keeping an orderly atmosphere in the school
		BCBG17F	6) Ensuring that there are clear rules for student behaviour
		BCBG17G	7) Addressing disruptive student behaviour
		BCBG17H	8) Creating a climate of trust among teachers
		BCBG17I	9) Initiating a discussion to help teachers who have problems in the classroom
		BCBG17J	10) Advising teachers who have questions or problems with their teaching
		BCBG17K	11) Visiting other schools or attending educational conferences for new ideas
		BCBG17L	12) Initiating educational projects or improvements
		BCBG17M	13) Participating in professional development activities specifically for school principals

Self-beliefs

Self-Beliefs	Item's number	Item description
Self-concept (like learning maths)	BSBM14A	a) I enjoy learning mathematics
	BSBM14B	b) I wish I did not have to study mathematics
	BSBM14C	c) Mathematics is boring
	BSBM14D	d) I learn many interesting things in mathematics
	BSBM14E	e) I like mathematics
	BSBM14F	f) It is important to do well in mathematics
Self-efficacy (confident in maths)	BSBM16A	a) I usually do well in mathematics
	BSBM16B	b) Mathematics is more difficult for me than for many of my classmates
	BSBM16C	c) Mathematics is not one of my strengths
	BSBM16D	d) I learn things quickly in mathematics
	BSBM16E	e) Mathematics makes me confused and nervous
	BSBM16F	f) I am good at working out difficult mathematics
	BSBM16G	g) My teacher thinks I can do well in mathematics with difficult materials
	BSBM16H	h) My teacher tells me I am good at mathematics
	BSBM16I	i) Mathematics is harder for me than any other subject

Appendix 7: example of school climate scale re-construction (winsteps)

Item fit

```
TABLE 10.1 bcgidnm5.sav                          ZOU684WS.TXT  Jul 30 11:20 2019
INPUT: 153 Person  5 Item  REPORTED: 153 Person  5 Item  9 CATS    WINSTEPS 3.73
-------------------------------------------------------------------------------
Person: REAL SEP.: 1.79  REL.: .76 ... Item: REAL SEP.: 5.51  REL.: .97

         Item STATISTICS:  MISFIT ORDER

-------------------------------------------------------------------------------------------
----------
|ENTRY   TOTAL  TOTAL           MODEL|   INFIT  |  OUTFIT  |PT-MEASURE |EXACT
MATCH|        |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%
EXP%| Item   |
|------------------------------------+----------+----------+-----------+-----------
+--------|
|     2    553    153   16.88     .44|1.50   3.8|1.82   5.7|A .63   .71| 72.5
64.2| BCBG11C|
|     5    574    153   15.50     .45|1.19   1.5|1.36   2.7|B .73   .71| 64.7
65.4| BCBG11H|
|     4    538    153   17.83     .43|1.15   1.3|1.16   1.4|C .80   .72| 59.5
63.1| BCBG11E|
|     3    657    153   10.30     .35| .83   -.9|1.07    .5|b .63   .65| 69.9
64.2| BCBG11D|
|     1    581    153   15.02     .45| .79  -1.8| .76  -2.0|a .71   .71| 71.9
65.5| BCBG11B|
|------------------------------------+----------+----------+-----------+-----------
+--------|
| MEAN   580.6  153.0   15.11     .43|1.09    .8|1.24   1.7|           | 67.7
64.4|        |
| S.D.    41.1     .0    2.60     .04| .26   2.0| .35   2.5|           |  5.0
.9|        |
-------------------------------------------------------------------------------------------
----------
```

Item-person map

```
TABLE 12.2 bcgidnm5.sav                          ZOU684WS.TXT  Jul 30 11:20 2019
INPUT: 153 Person  5 Item  REPORTED: 153 Person  5 Item  9 CATS    WINSTEPS 3.73
-------------------------------------------------------------------------------

          Person - MAP - Item
              <more>|<rare>
   22            .  +
   21          .### T+
   20               +T
   19            .#  +
   18               +S BCBG11E
   17           ###  +  BCBG11C
   16              S+
   15               +M BCBG11B  BCBG11H
   14         .####  +
   13               +S
   12     #########  +
   11               +
   10        ###### M+T BCBG11D
    9               +
    8       .######  +
    7               +
    6       .######  +
    5               +
    4         .#### S+
    3               +
    2           .##  +
    1            .#  +
    0               +
   -1              T+
   -2            #  +
              <less>|<frequ>
 EACH "#" IS 3. EACH "." IS 1 TO 2

↟TABLE 12.5 bcgidnm5.sav                          ZOU684WS.TXT  Jul 30 11:20 2019
INPUT: 153 Person  5 Item  REPORTED: 153 Person  5 Item  9 CATS    WINSTEPS 3.73
-------------------------------------------------------------------------------
```

Unidimentionality

```
TABLE 23.0 bcgidnm5.sav                          ZOU684WS.TXT  Jul 30 11:20 2019
INPUT: 153 Person  5 Item  REPORTED: 153 Person  5 Item  9 CATS    WINSTEPS 3.73
-------------------------------------------------------------------------------

     Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                 -- Empirical --    Modeled
Total raw variance in observations     =         10.3 100.0%          100.0%
  Raw variance explained by measures   =          5.3  51.6%           54.7%
    Raw variance explained by persons  =          4.6  44.3%           47.1%
    Raw Variance explained by items    =           .7   7.2%            7.6%
  Raw unexplained variance (total)     =          5.0  48.4% 100.0%   45.3%
    Unexplned variance in 1st contrast =          1.8  17.7%  36.6%
    Unexplned variance in 2nd contrast =          1.5  14.5%  29.9%
    Unexplned variance in 3rd contrast =          1.0   9.8%  20.3%
    Unexplned variance in 4th contrast =           .6   6.2%  12.7%
    Unexplned variance in 5th contrast =           .0    .3%    .5%

       STANDARDIZED RESIDUAL VARIANCE SCREE PLOT

        VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+--+
   100%+  T                           +
       |                              |
  V 63%+                              +
  A    |     M                        |
  R 40%+        P     U               +
  I    |                              |
  A 25%+                              +
  N    |                              |
  C 16%+                  1           +
  E    |                     2        |
    10%+                              +
  L    |                        3     |
  O  6%+           I                  +
  G    |                           4  |
  |  4%+                              +
  S    |                              |
  C  3%+                              +
  A    |                              |
  L  2%+                              +
  E    |                              |
  D  1%+                              +
       |                              |
   0.5%+                           5  +
       +--+--+--+--+--+--+--+--+--+--+
          TV MV PV IV UV U1 U2 U3 U4 U5
         VARIANCE COMPONENTS
```

Appendix 8: Qualitative data gathering guidance

QUALITATIVE DATA COLLECTION INSTRUMENTS

Multiple data sources will be conducted to explore the climate of a school and how this climate affect student learning. Sources including interviews with head teachers and focus group interviews of teachers and students.

Interview

Head teacher will be interviewed in their offices for approximately sixty minutes during a formal visitation. An interview protocol will be administered as guidance and can be flexible. The focus of the questions is on the experience of head teachers in creating their school climate, how to do evaluation of the climate, and how this climate affect their student learning. Informal follow-up interviews and visitation will be conducted to deepen the researchers understanding on an as-needed basis.

Before starting the interview, the researcher give an explanation about school climate and each aspect that constructed school climate. The purpose of this explanation is to have the same understanding of school climate and its aspects, so the interview going to more focus. Due to strict selection of the school that will participate in the study, if the head teacher is not possible to do interview in any reasons, the interview will be conducted to other significant person in the school who responsible to manage the school.

Interview Guidance

The interview protocol is based on dimension of dynamic model of educational effectiveness, particularly on how to do evaluation of the school learning environment (Creemers & Kyriakides, 2008, p. 135). The model give explanation how to asses of each dimension and has been adapted to fit with school climate aspect as recently reviewed by Thapa et al. (2013).

General question

1. How do you perceive of school climate of the school? Could you explain?

Question for each dimension of educational effectiveness

Educational effectiveness dimensions	School climate aspects	Questions
Frequency It is assessed by discovering 1) how often school collects evaluation data and 2) how many sources were used	Safety Relationship Teaching and learning Institutional environment Improvement process	1. Do you have a (school climate aspect) policies? How often do you evaluate this policy?
	General	2. How many sources of evaluation data were used for all aspects of these school climate?
Focus It refers to which aspects of school climate (safety, relationship, teaching and learning, institutional environment) are assessed.	Safety Relationship Teaching and learning Institutional environment Improvement process	1. What is the focus/concern of this school in creating school climate? Why? (for example, the school focused on creating physical aspect of school, because they have enough resources to do so) 2. How school gathered information about this focus? Is it from a specific evaluation (the toilet is not clean enough) or from a general evaluation (student do not like the school building)?
	General	3. What is the purpose of collecting such evaluation (for formative or for summative reason
Stage It is assessed by looking at the period in which data are gathered.	General	1. When did the school climate evaluation take place? Every semester, end of the year, or as a continuous process? 2. How did school review their evaluation mechanism? Did school adapt the evaluation mechanism in order to collect relevant and proper data at each stage?
Quality Quality is evaluated by looking at the validity and reliability of instrument used in collecting school climate data.	Safety Relationship Teaching and learning Institutional environment Improvement process	1. What instrument do school use to collect school climate data? Why? Is the instrument valid and reliable? (Survey, interview, etc.)
	General	2. Is the information collected from evaluation is used in making decision in order to improve school climate?
Differentiation This dimension discusses to the degree to which the school leaders give more emphasis on conducting evaluation for specific aspect of school climate? For example, if the lack of school resources is a focus of the school. The school might	Safety Relationship Teaching and learning Institutional environment Improvement process	1. Which aspect of school climate that school give more emphasis at the time? Why?

choose to gather data more often on this aspect than any other aspects of school climate.		
	General	2. How does this emphasis help school to solve that problem?

Focus group interview (FGI)

Teachers and students will be arranged in different focus group sessions. The group is about 5 - 7 participants. It will take their time for about forty-five to sixty minutes. To make this discussion on the track, the moderator will lead the discussion based on guiding questions, however it is also a room for flexibility. The guidance focusing on how teacher and student feel or perceive their school climate and how those perception or feeling affected them in the teaching learning process. The guidance is based on school climate aspect as reviewed by Thapa et al. (2013).
The researcher will do every effort to conduct the session well organised, however if the sessions can't be conducted, the researcher will use another way to gather teachers and student perception of their school. It can be done by interviewing some teachers and students.

FGI Guidance

Student

General question

1. Can anyone tell me about the school? What is the best part of the school? And also what need to improve? (opening question)
2. Which aspect of school climate is the most important on your academic and personal development? (closing question)

Questions based on school climate factor

School climate aspects	Questions
Safety	1. Are you feeling safe in the school? Why? 2. Have you ever been bullied or saw other students was being bullied? How did school respond on the issue? 3. Do you think the school has a clear rule in term of safety? (**frequency)** 4. Do you think school evaluate the rules in regularly? How? **(frequency)** 5. Do you think that school has specific issues in safety which is the main target to create safety school? What? **(focus)** for example: to minimise bullying 6. What do you think of other purpose in creating safety climate in this school? **(focus)** for example: to build positive relationship, or may be build discipline, minimise absenteeism. 7. Do you think that the rules applied has been used for long period or just been applied? Do school consistently apply the rules? **(stage)** 8. What do you think about the quality of safety in this school? Do school take the safety issue seriously? **(quality)**

	9. Do school have more focus on specific group on implementing safety rules? For example: school more focus on younger student or minority student. **(differentiation)**
Teaching and Learning	1. Do school/teacher give support to student achievement and or personal development? How? **(frequency)** 2. How do you feel you about your classroom? Do you find your teacher easy to understand? Do they make you interest in their lesson? **(frequency)** 3. How do teachers/head teacher support your learning? Do you have chance to get personal support? **(frequency)** 4. What is the focus of that support? It is giving more emphasis on to improve your achievement. Or to improve your personal development, or others? **(focus)** 5. Do you think that their support take place a long time ago or just been applied? Why? **(stage)** 6. Do you find that how the teacher gives their lesson encourage you to do more in your school? **(quality)** 7. Do teacher support only on specific group? For example, for only low achievement student or vice versa **(differentiation)**
Interpersonal Relationship	1. How do you feel about relationships between students? **(frequency)** 2. What is the focus of interpersonal relationship that been built in your school? Can you explain? **(focus)** 3. What are other purposes in building relationship? **(focus)** 4. How long is the relationship exist? **(stage)** 5. Does the relationship support your academic and personal development? How? **(quality)** 6. How do you treat other students from minorities' background? Is there any differentiation? (Ethnic, religion, SES, etc.) **(differentiation)** 7. How do you feel about teacher-student relationships? **(frequency)** 8. What is the focus of the relationship? For improving academic performance. For creating a warmth climate. Others? **(focus)** 9. Do you think that kind of relationship have been done for long time or only in recent day? **(stage)** 10. Does the teacher encourage you to do your best in academic or personal development? Are they supportive enough? **(quality)** 11. Is the relationship building more emphasis on specifics group of students? **(differentiation)** 12. Does your parent support your learning? In what why? **(frequency)** 13. Do they ask you about school? (frequency) 14. What did they do to support your learning? **(quality)** 15. Do you discuss what you did in the school? **(quality)**
Institutional environment	1. What do you like about your school? Why? 2. Do you like being in the school? Why? **(frequency)** 3. Do you involve in a co-curricular or after school program? Why? **(frequency)** 4. **What is the focus of the school in creating the school physical environment? (focus)** 5. Is there any other purpose? E.g., increase safety, building pride, etc. **(focus)** 6. How long do you feel that you being like in the school? **(stage)** 7. Do you think that you school environment gives an effect to your academic and personal development? **(quality)** 8. Do you think this school has enough facilities to support your learning? (Classrooms, library, laboratories, canteen, etc.? **(quality)** 9. How about the school building, is it good enough? (Appearance, cleanliness, etc.?) **(quality)**

	10. Do you feel that this school only suit for certain group of students? **(differentiation)**

7.7.1.1 Teacher

7.7.1.1.1 General question

1. Can anyone tell me about the school? What is the best part of the school? And also, what need to improve?

Questions based on school climate factor

School climate aspect	Questions
Safety	1. Are you feeling safe in the school? Why? 2. Do you think the school has a clear rule? Do school consistently apply the rules? 3. Have you ever been bullied or saw other students was being bullied? How did school respond on the issue?
Teaching and Learning	1. How do school/head teacher give support to student achievement and or personal development? 2. How do teachers/head teacher support student learning? Do you give personal support to student? 3. What do you do to attract student in your lesson? 4. Do you have special approach to make student understand of your lesson? 5. How do you feel you about your classroom? Do you think your student involved in your lesson? 6. How do you expect of your student achievement?
Interpersonal Relationship	1. How do you feel about relationships between students? Are they behave well, respect each other? 2. How do you treat student from minorities' background? (Ethnics, religions, SES, etc.) 3. How do you feel about teacher-student relationships? 4. How do you develop relationship with your students? 5. Do you encourage your students to do their best both in academic and personal development? Are they supportive enough? 6. How do you support student to improve their academic and personal development? 7. How do you feel about your professional relationship with other teachers? And how about personal relationship with other teachers?
Institutional environment	1. What do you like about your school? Why? 2. Do you like being in the school? Why? 3. Do you satisfy enough with your career as a teacher? 4. How do you compare your enthusiasm right now and at first time you were appointed as a teacher? 5. Do you think this school has enough facilities to support teaching-learning activities? (Classrooms, library, laboratories, canteen, etc.?) 6. How about the school building, is it good enough? (Appearance, cleanliness, etc.?)

Appendix 9: New themes

Cited Words	Code 1	Code +effectiveness dimension	Final code
Actually, that's the reason we want to return to the central core *of madrasah*. Since the implementation of the National Exam, *madrasah* has almost lost its identity, because all the energy is focused on the National Exam, especially for the final year of study (MORA policymakers).	The focus on exam	Curriculum compliance: focus	Orientation of education
Madrasah must include general curriculum lessons, so, the student must take 100% general curriculum with the addition of religious subjects mentioned earlier (*fiqh, Quran studies,* etc.) 2 hours/week. Means that our children learn more subjects compared to their friends in general schools and as a consequence. (MORA policy maker)	Curriculum employment	Curriculum compliance: frequency	Orientation of education
To get a better pass rate in the National Exam, we did the exam like try out at least once before the student facing the real national exam (Local policymaker).	Curriculum demand	Curriculum compliance: frequency	Orientation of education
Yes, that's the dilemma, we want the children to understand, but the curriculum mandates that the material must be finished. So, there is a lot of subject matter to be delivered, but time is limited. Because of the objectives of this curriculum, we cannot implement the parallel methods that we have compiled, so we only give a quick formula, questions, in order to pass the examination, that's it. Because of curriculum wants us to do that. (Teacher, *Madrasah* B)	Focus on finishing curriculum demand	Curriculum compliance: focus	Orientation of education
So, the material tested in the National Exam for example, so that is taught just that, others do not, as a result, the child is trapped in a	Focusing on the national exam	Success in the national exam: focus	Orientation of education

pattern of evaluation system early (MOEC policymaker)			
We also have a program for tackling the national exam. In almost every school, they have special programs to face the National Exam, for example, by giving additional lesson time focusing on how to answer the national exam tests. (Local policy maker).	Focusing national exam	Success in the national exam: focus	Orientation of education
The national exam becomes a focus of the local government in the region because if more school failed in exam or low pass rate, it would affect the leader electability rate in the area. For example, most of governor or mayor stressed that the school must reach the target of 100% pass rate (MOEC policymaker).	Political agenda	Exam orientation: frequency	Orientation of education
We have chosen *madrasah* because it gives us more religious lessons, strong prayer discipline, Arabic lessons, as well as Islamic moral lessons (student, *Madrasah* B)	Promoting Islamic values	Culture: focus	Culture
In our city (with a large proportion of Muslim population), by default, the common demand that, when the students graduated from elementary school, they must be good at reading and writing the Qur'an (Local policymakers)	Promoting Islamic values	Culture: frequency	Culture
So, in the development of school culture is much more influenced by cultural and religious values (Local policymakers)	Cultural and religious value	Culture: focus	Culture